ABSTRACT

Title of Dissertation: STUDENT LEARNING OUTCOMES AND THE IMPACT ON PRINCIPALS' SELF-EFFICACY BELIEFS

James Scott Smith, Ph.D., 2017

Dissertation Directed by: Patricia M. Richardson, Ph.D.
Department of Education

Student Learning Outcomes (SLOs) as defined by the Maryland State Department of Education are "a specific, rigorous, long-term goal for groups of students that educators distinguish to guide instructional and administrative efforts." They serve "as one of the measures of student growth for the State Principal Evaluation Model and represent a significant portion of the evaluation" (MSDE, 2013, p. 15). While much work has been done in Maryland with regard to principal evaluation forms, processes, and timelines, evaluators are given limited guidance on how to work with principals to attain their SLOs. A deeper understanding of social cognitive theory—specifically self-efficacy beliefs—may help inform, organize, and enhance the SLO process. An emerging field of study may provide meaningful insight into the work done with leaders to set purposeful SLOs and to support them through the process.

This directed study of 26 principals contributed to the research about principal self-efficacy beliefs and gathered principals' reactions to their evaluation experiences relative to the SLO process. The study used the Principal Self-Efficacy Survey (PSES) followed by a series of questions eliciting reactions to the SLO process from the

principals who have gone through an evaluation cycle. The specific questions asked of principals, relative to their SLO process, were aligned with Albert Bandura's four determinants of self-efficacy: mastery learning, vicarious learning, verbal persuasion, and emotional arousal. The study confirmed higher mean self-efficacy scores by principals who met their SLOs and reinforced the determinants of Albert Banduras were present in the process used by this school system. By learning more about principal self-efficacy beliefs and factors that positively or negatively influence such beliefs, recommendations can be made for tailoring the SLO process to increase principal self-efficacy beliefs.

STUDENT LEARNING OUTCOMES AND THE IMPACT ON PRINCIPALS' SELF-EFFICACY BELIEFS

by:

James Scott Smith

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Education
2017

Advisory Committee:

Professor Patricia M. Richardson, Chair
Professor Olivia Saracho, Dean's Representative
Professor Margaret J. McLaughlin
Professor Helene Cohen
Professor Donna Wiseman

Dedication

I dedicate this dissertation to my family and my professional colleagues who have supported me throughout this process. To my wife, Michelle, who constantly believed in the value of an advanced degree and was willing for me to sacrifice time with her and my three boys, I will be forever thankful. I am also incredibly grateful to all my colleagues, who provided encouragement and perspective throughout this process.

Acknowledgements

I am particularly indebted to my advisor, Dr. Patricia Richardson, who was always available for a phone call, impromptu meeting, or extended discussion of what leadership should really look like. She understands balance and the need to give every voice a compassionate ear. Dr. Richardson provided me my first leadership opportunity and always offered me a steady hand as I progressed in the school system. I have always admired her as a leader and have sought to emulate her. I am grateful to Dr. Margaret McLaughlin, who set the high standards for this cohort and accepted nothing less. Finally, I thank my professors in the program, who brought to class a wide range of leadership experience that they were always willing to share with the group. They offered a perspective of education beyond grades K–12, enabling members of the cohort to see the larger picture of the forces at work in public education and opening our eyes to how all the pieces of this machine we call education works.

Table of Contents

List of Figures and Tables

Section I: Introduction

Introduction of the Problem

Federal regulations, including No Child Left Behind, Race to the Top (RTTT), and most recently, Every Student Succeeds Act (ESSA), are all predicated on evaluating school performance by implementing an annual accountability system. The accountability is spread across students, teachers, and administrators. While student assessments and teacher evaluation systems have been at the center of most discussions, annual principal evaluation systems are also required and must be tethered to student achievement results.

In 2014, the Maryland State Department of Education (MSDE) published their model for principal evaluation, "The Maryland Teacher and Principal Evaluation Guidebook." One half of the model was based on qualitative measures that had been in use for the previous five years, but the other half was based on new quantitative measures, the achievement of Student Learning Outcomes (SLOs) (p. 7). SLOs, as defined in the Guidebook, are "specific, rigorous, long-term goal(s) for groups of students that educators distinguish to guide instructional and administrative efforts" (p. 15). According to the Guidebook, principals meet with their evaluator at the outset of the year, set performance goals on measurable student outcomes, monitor those outcomes throughout the year, and conclude their year with half of their annual evaluation determined by the attainment of their SLOs. While a process and formula have been established for this evaluation model, there is need for more work with evaluators of principals to set meaningful, authentic, and rigorous goals and, once set, "provide meaningful feedback and dialogue around their instructional leadership" (Slotnik, Bugler,

& Liang, 2015, p. 3).

An emerging field of study may have meaningful insight into the work done with leaders to set purposeful SLOs and to support them through the process. A principal's belief that he or she can successfully handle a specific, future challenge plays an important role in the principal's actual success (Tschannen-Moran & Gareis, 2005). According to Tschannen, "principals with a strong sense of self-efficacy have been found to be more persistent in pursuing their goals but were also more flexible and willing to adapt their strategies based on contextual conditions" (p. 5). When one considers principal evaluation and the inclusion of targeted teacher and student outcomes, principals' self-efficacy should be considered. Tschannen states that a principal's self-efficacy belief "exhibits compelling ramifications when considering the central function of the leadership of others that principals fulfill. Specifically, the leadership behavior of principals has been linked to teacher efficacy beliefs" (p. 6).

While much work has been done in Maryland with regard to principal evaluation forms, processes, and timelines, evaluators receive limited guidance on how they are to work with principals to attain their SLOs. A deeper understanding of social cognitive theory—specifically, self-efficacy beliefs—may help inform, organize, and enhance the SLO process.

Principal Evaluation in the 21st Century

In the era of big data and performance accountability, the measurement of principal effectiveness has been the focus of several studies. Seminal among them in the last decade is the work of the Wallace Foundation—a national philanthropic organization

dedicated to strengthening educational leadership. In March 2009, the Wallace Foundation published "Assessing the Effectiveness of School Leaders: New Directions and New Processes." The report stated, "effective leadership is vital to the success of a school. Research and practice confirm that there is slim chance of creating and sustaining high-quality learning environments without a skilled and committed leader to help shape teaching and learning" (Wallace, 2009, p. 1). The report further posits a "shift from building managers to learning leaders first and foremost is well documented, and is further backed by research indicating that leadership is second only to teaching among school-based factors in influencing learning" (p. 1). The paper's central argument is "that leader assessment is an important but largely under-developed part of the web of policies, practices and incentives needed to support our nation's principals. When used to enhance performance as well as ensure accountability, assessment can be a driving factor in helping leaders develop the behaviors and skills that will improve learning for all students" (p. 11). Wallace captures the central belief driving RTTT and ESSA, that evaluation tools must be directly driven by quantifiable student achievement data. The Maryland Teacher and Principal Evaluation Guidebook specifically states this requirement of "student growth" as a significant component of principal evaluation as mandated by RTTT (p. 4).

Elements of Effective Principal Evaluation

According to a report coauthored by the National Association of Elementary School Principals (NAESP) and the National Association of Secondary School Principals (NASSP), effective principal evaluation systems are based on seven beliefs:

1. Created by and for principals;

2. Part of a comprehensive system of support and professional development;
3. Flexible enough to accommodate differences in principals' experiences;
4. Relevant to the improvement of principals' dynamic work;
5. Based on accurate, valid and reliable information gathered through multiple measures;
6. Fair in placing a priority on outcomes that principals can control; and
7. Useful for informing principals' learning and progress (Clifford, M., Ross, S., 2013).

This same disposition is seen in the "Tips and Tools for Evaluating School Principals" created by the National Comprehensive Center for Teacher Quality, which states that evaluation systems must clearly establish goals and expectations, use valid and reliable data that help inform professional development, be research based, and include multiple forms of assessment to yield a holistic view of principal performance (Browns-Simms, 2010).

Principal Evaluation as a Process

Also supporting these tenets is a phenomenological study done in 2012 by Oksana Parylo that explored 16 principals' life experiences of being evaluated and had many of the same findings (Parylo, 2012). Principals interviewed admitted that their "roles and responsibilities have gone through considerable changes in the accountability era" and that those who had been principals for more than ten years had seen their evaluation evolve from "more of a managerial type, where you were evaluated on how well you managed people, dealt with a budget and dealt with the community" (p. 223). Evaluation systems now are more data driven and performance based. To balance this, principals valued evaluation systems that were a process—not an event—built on transparency,

valuing dialogue, predicated on trust and respect, and underpinned with feedback and support (Parylo, 2012).

Trust and Support

Trust and support was also a recurring theme in the research done by the National Comprehensive Center for Teacher Quality (Clifford, M., 2012). The report concluded that new evaluation systems should not only hold principals accountable for performance, but must also generate "trust among stakeholders" and "support principals' continued growth" (p. 59).

This has not always been the case. According to research conducted in 2012 by the NAESP and the NASSP, "Many principal evaluation instruments are neither technically sound nor useful in improving principal evaluation" (Clifford, M., 2013, p. 2). To counter this claim, much work has been done to make principal evaluation instruments more empirically based by incorporating value-added models, which has met with multiple obstacles (Lipscomb, 2012). Mathematica Policy Research presented findings for the piloting of the statewide Pennsylvania principal evaluation system and found that factors beyond the principal's control made isolating pure principal effects daunting. They concluded that the best model was one where a principal is compared to a previous principal at the same school because that scenario allows for control of variables. As this is not a realistic model, they instead developed an average effectiveness of value-added during an analysis period (Lipscomb, 2012). After modeling and compiling all the data, the results showed "moderate consistency"—even though they captured the contributions of the entire school—not just those factors directly under the principal's control (Lipscomb, 2012). While this approach addressed the "technically

sound" aspect of evaluation by providing a statistical model, it does not align to the belief espoused by the NAESP or NASSP and does not inform a principal's learning or professional development.

Reliable Instruments

Developing technically sound principal evaluation instruments was the basis for the development of the Vanderbilt Assessment of Leadership in Education (VAL-ED), which stated that "principals are often evaluated with the use of instruments with no theoretical background and little, if any, documented psychometric properties" (Porter, A., 2010, p. 135). The case study sought to test an education leadership performance assessment system for measuring the effectiveness of principal leadership behaviors in a 360-degree program that took input from the teacher, principal, and principal supervisor. The study revealed that the iterative process of revise, test, revise, and test was an essential element in creating a psychometrically valid and reliable tool (Porter, A., 2010). This work emphasizes the importance of all stakeholders being part of the process and having a voice in its development and refinement.

Goal-Driven Evaluation

At the core, the most effective principal evaluation systems are those that have been created with direct input from the principals and reflect the reality of their daily job (Catano, 2006). Such evaluations explicitly state the expectations and the manner in which the principal will be measured (Clifford, 2012; Lipscomb, 2012; Parylo, 2012). They are flexible tools that evolve as the role of the principal responds to the needs of the school and staff (Casserly, 2013). Principal evaluation systems must be relevant to the

work being done by the principal and must incorporate the standards of practice that are widely accepted (Browns-Simms, 2010). The goal of evaluation is to improve performance, and this is done by selecting relevant measures that reflect the reality of a principal's job (Parylo, 2012), (Clifford, M., 2013), (Clifford M., 2012). Principals have a complex set of job responsibilities that require finesse and flexibility (Catano, 2006). Evaluation instruments must inform professional development, as the cycle of evaluation is the setting of expectations and goals (Canole, 2013). Evaluation systems must be transparent and systematically applied to all principals with the highest priority placed on outcomes the principals control (Clifford, M., 2013). There must be no "surprises"; if there is an area of need, it must be identified in a manner that can be addressed and corrected (Parylo, 2012). Considerable evidence shows that principal evaluation must be collaboratively developed, reflective of the work of the principal, and target driven.

Principal Evaluation in Maryland

In August 2010, the Maryland Department of Education (MSDE) was awarded a $250,000,000 RTTT grant. For those funds, MSDE agreed to four assurances, one of which was "Great Teachers and Leaders." The "Great Teachers and Leaders" central assurance was that 50% of all principal and teacher evaluations would be based on student growth—30% of which would be state approved and 20% approved locally. Over the past seven years, the state has developed the student growth element of evaluation into SLOs. SLOs as defined by the MSDE are "a specific, rigorous, long-term goal for groups of students that educators distinguish to guide instructional and administrative efforts." They serve "as one of the measures of student growth for the State Principal Evaluation Model and may represent 20% - 35% of the evaluation" (MSDE, 2013, p. 15).

Education Reform Act of 2010

As the MSDE was applying for the RTTT grant, Governor Martin O'Malley was signing into law the "Education Reform Act of 2010." The act required teacher and principal evaluations to include student growth as a "significant factor" in the overall evaluation (Education Reform Act of 2010). In its original iteration, it required student growth to be 50%, but subsequent revisions have it now defined as "a significant component of the evaluation and as one of multiple measures" (Education Reform Act of 2010).

Teacher or Principal Evaluation Field Test

Megan Dolan, Ed.D., was commissioned by MSDE to do a Statewide Summary Report of the Evaluation Field Test in 2013 by conducting 66 focus groups in 22 Local Education Agencies (LEAs) participating in the Teacher and Principal Evaluation (TPE) Field Test. In each county, "the Field Test Monitor (FTM) met separately with three focus groups consisting of central office personnel, school level administrators (including principals, assistant principals, and content supervisors), and teachers" (Dolan, 2013, p. 2). Her findings indicated that "stakeholders in all of the LEAs are working very hard to establish teacher and principal evaluation systems that will be ready for full implementation starting in August, 2013" (p. 5). She also noted several areas of challenge, which centered on the achievement data for SLOs. They included, "an apparent misalignment between the availability of student assessment scores and the schedule for making personnel decisions that are dependent on these scores. This prompted some LEAs to rely on a mixture of professional practice data from the current school year and the results of student scores from the previous school year" (p. 7). This

was exacerbated by the "confluence of multiple new initiatives that require extensive investment of resources at all levels, including the transition to the Maryland Common Core Curricula and the anticipation of the new assessments from the Partnership for Assessment of Readiness for College and Careers (PARCC)" (p. 6). One of the final challenges she stated was "the emphasis on teacher evaluation rather than principal evaluation. As it also did in many other states, this led some LEAs to devote the bulk of their resources to developing and implementing teacher evaluation, while far less attention was paid to principal evaluation, and some of those models are not yet complete" (p. 7).

The Maryland Teacher and Principal Evaluation Guidebook

In September 2013, MSDE released a revised version of "The Maryland Teacher and Principal Evaluation Guidebook." The document began with an overview of the four mandates moving TPE forward in Maryland. They were: The Education Reform Act of 2010, Elementary and Secondary Education Act (ESEA) Flexibility Waiver, COMAR Title 13A.07.09, and the Maryland RTTT Grant Application (MSDE, 2013). Under the Education Reform Act of 2010 school systems were required to include "data on student growth as a significant component of the evaluation and one of multiple measures" (p. 40). Under the ESEA Flexibility Waiver, school systems must include Maryland School Assessment scores to account for 20% of a principal evaluation—to be done specifically through "one Student Learning Outcome (SLO) with a data point from statewide High School Assessments (HSAs) in the evaluation" (p. 4). The final mandate for principal evaluation was under the RTTT, requiring that "the evaluation rating reflect professional practice as 50 percent of the value and student growth as 50 percent of the value" (p. 5).

MSDE provided a framework for development of the principal SLO, detailing sequential tasks for the principal evaluation cycle: from the initial conference, implementation and observation, evaluation, and finally data review to inform next steps (MSDE, 2103). Also included in the SLO portion of principal evaluation was the School Progress Index, defined as a tool that evaluates schools on a continuous scale based on the variables "predicated on a series of local annual measureable objectives which examine achievement, gap, and growth in elementary and middle school, and college and career readiness in high school" (p. 13).

MSDE created and provided to LEAs a template for the creation of principal SLOs that included: Objective Summary Statement, Data, Student Population, Learning Content, Instructional Interval, Target, Evidence of Growth, Leadership and Professional Development (MSDE, 2013). Coupled with this was a rubric for evaluating the identified SLOs that detailed four elements: Priority of Standard, Rigor of Target, Quality of Measure & Evidence, and Action Plan, with criteria for each (MSDE, 2013). For each element, the principal evaluator is to determine the level of SLO attainment: full, partial, or insufficient (MSDE, 2013).

Real Progress in Maryland

In September 2014, the Mid-Atlantic Comprehensive Center released "Real Progress in Maryland, Student Learning Objectives and Teacher and Principal Evaluation." For the report "researchers collected both qualitative and quantitative data from educators across the state, including (a) interviews of superintendents, teacher association leaders, central administrators, and principals from 17 districts; (b) focus groups of teachers from four districts; and (c) a statewide survey of teachers, principals

and other educators. The final survey analysis is based on the responses of 16,314 educators from 23 local school districts, with district response rates ranging from 5.7% to 72.6% and a collective 23 district response rate of 31.3%" (Slotnik, Bugler, & Liang, 2014, p. 1). From this report, "principals have a more positive impression than teachers that SLOs are being implemented effectively, that data are used to develop SLOs, that there are opportunities to confer with principals about SLOs, that there are opportunities to engage in a mid-year review, and that evaluation results will be used to inform professional development plans" (p. 3). The report focused on the teacher-created SLOs and how the principals manage this process, to include the need to "develop an expanded series of principal leadership prompts" and "provide a specific stream of support customized to high school principals and their immediate supervisors" (p. 4). Actual review of principal SLOs is not included in the report. It focuses instead on the principal's interaction with teachers as they worked through their SLOs. This was brought to light when the study was repeated the following year.

Change in Practice in Maryland

In September 2015, the Mid-Atlantic Comprehensive Center released "Change in Practice in Maryland, Student Learning Objectives and Teacher and Principal Evaluation." The data for this study included interviews from twelve districts and statewide surveys of teacher, principals, and other educators from all twenty-four districts. Of note was educator anxiety about the concurrent rollout of the new evaluation systems, new standards, curricula, and assessments—all of which require substantial input of resources and time (Slotnik, Bugler, & Liang, 2015). Principals specifically reported that they "want to see districts pay more attention to principal evaluation, giving

principals the opportunity for meaningful feedback and dialogue around their instructional leadership" (p. 3). Principals also stated that "unmet needs in key capacity building areas affect their ability to support and guide their respective school staffs in the same areas" (p. 3). The report later revealed that "while the statewide focus is on both teacher and principal evaluation, a continuing trend in principals' comments is that the principal evaluation component is receiving less attention than the teacher component" (p. 52). Specific quotes from principals in this report highlighted this frustration:

> From the principal evaluation perspective, it has really been an afterthought for our system. We have gotten no professional development. We were given the rubric and told this is how we will be assessed, but no discussion on what it means. It's been a source of frustration to principals.
>
> –*Principal* (p. 53)
>
> It hasn't helped as a principal yet, but as I look at the rubric and expectation, I think it can be. It's definitely allowed me to reflect and that's going to make me grow. There just needs to be more professional development.
>
> –*Principal* (p. 53)
>
> I don't have one thing that's really meaningful. It's just me telling them how great I am, and they agree. I never leave with something meaningful to improve on or a growth piece and that would be nice. There are many things I can improve on, and I don't get that.
>
> –*Principal* (p. 53)
>
> We are supposed to have a conference with the assistant superintendent of instruction – we set goals at the beginning of the year… We rate ourselves and at

the end of the year we do a narrative based on how well we met those goals and then we are supposed to have a conference with the Superintendent but that hasn't happened in 3 years.

–*Principal* (p. 53)

I don't think it was that great of feedback to be honest with you. I didn't feel like I got a lot of specific things that I could take from it and go back and work on. I felt that way about the previous evaluation model too. I didn't feel like it was specific enough to give you something concrete you can do and improve.

–*Principal* (p. 53)

Stay the Course

The final study done by the Community Training and Assistance Center, "Stay the Course, Teacher and Principal Evaluation in Maryland," was published in September 2016. The study drew from 21,916 survey responses from principals and teachers and extensive interviews of educators at central and school site levels (Slotnik, Bugler, & Liang, 2016). The report concluded that to make Teacher and Principal Evaluation manageable requires the creation of "a meaningful principal evaluation system to better understand and address the needs of principals" (p. 3). Their key findings were that "principals and teachers still need support" (p. 11). Fifty-seven percent of responding principals stated that they need more support in the form of "receiving feedback from district administrators" and 69% cited the need for more support "using data from SLOs and observations to make improvements in (their) instructional supervision" (p. 17). One principal stated that "the pressure is on teachers and principals. Where is the pressure on the central office to support me and my colleagues?" (p. 35) The report stated that

"principal leadership is key to the success of any serious initiative at the school level" and need exists for "a principal evaluation system that can strengthen their performance and contribute to their growth as school leaders" through more "coaching and formalized supports as they implement new evaluation systems" (p. 49).

Principal Evaluation in the School System

The school system is located at the southernmost tip of Maryland, 60 miles south of the nation's capital. St. Mary's County, traditionally known for its agricultural and fishing communities, is now best known for its world-class technology corridor that supports the operations of the Patuxent Naval Air Station. With over 18,000 students and 2,000 staff members, SMCPS is large enough to offer a wide variety of academic programs and services and small enough to maintain an environment of meaningful relationships, helpfulness, and personalized attention. Students have the opportunity to learn through a rigorous comprehensive program of instruction or to choose one of the Educational Pathways available, including a Public Charter School; Fairlead Academy; Tech Connect; the Academy of Finance; the Academy of Global and International Studies; and the Science, Technology, Engineering, and Mathematics Academy (www.smcps.org).

The school system employs 25 principals, 3 academic deans, and 41 assistant principals. MSDE requires that each educational leader be evaluated annually and that the evaluation reflect qualitative professional practice measures and quantitative student growth measures.

Leadership Performance Assessment System

The most recent version of the school system's Leadership Performance Evaluation System (LPAS) now mirrors the state model—with half devoted to professional practice and half devoted to SLOs (Leadership Performance, 2014). While the professional practice portion is clearly defined, the SLO portion is a work in progress and still a relatively new concept. The problem faced by the school system is how to implement this mandate and support the principals' focus on their continued improvement as educational leaders in the system. While MSDE has provided a template, meeting schedule, and evaluation rubric, there has not been the time or study done to determine the quality of principal SLOs and the impact they may have on a principal's perception of the tool and the value they place on the process. While Maryland is leading this work, searching for relevant, recent research on principal SLOs yields nothing more that the summary guidance published by MSDE in 2014. Moreover, no study has investigated whether including SLOs in the evaluation process affects principals' perception of the process or the tool.

Student Learning Outcomes

The school system is compelled to comply with the MSDE directives and has committed itself to the purposeful evaluation of principals, with 50% of the model predicated on professional practice and the other 50% based on SLOs collaboratively developed between the principal and the evaluator (SMCPS, 2015).

Over the past four years, one of the most challenging elements of this work is establishing meaningful SLOs that capture the true work of principals and honors their efforts. While the professional practice portion is drawn from a list of standards with each

being defined by accepted rubrics, the SLO portion varies from leader to leader, as principals have the latitude to set individual, specific goals across a broad spectrum (SMCPS, 2015). One SLO must address a student achievement target and the other must address a culture or climate initiative (SMCPS, 2015). The 2015 LPAS manual directs that "the SLOs will be aligned to measures of student learning that are the focus of school improvement. For example, a school may set SLOs based on school-wide performance of SATs or AP exams, or the SLOs can be based on another measure the school is working toward (e.g., graduation rate)" (p. 6.2). The manual also chronicles the following process for setting performance targets:

1. Leaders should review school or content data to determine current levels of performance that include proficiency levels and area(s) of focus. Data considered may include pre-assessment data given that fall, as well as data carried forward from previous school years.
2. Complexity factors will be jointly discussed by the evaluator and employee, and considered by the evaluator.
3. The leader reviews the performance targets set at the system level, as well as school improvement targets to identify the School SLOs.
4. The leader should identify at least two (2) SLOs for the current school year - one focused on achievement and one focused on leadership.
5. If applicable, the leader selects a content and focus group for which the learning evidence will be used. This is the attributable group.
6. The leader meets with the evaluator to review the School SLOs.
7. Considering complexity factors that may impact overall performance, the

> evaluator and leader confer, and the evaluator approves the specific, measurable School SLOs. Both the leader and evaluator sign documentation of these targets by October 30 (SMCPS, 2015, p. 6.2).

One SLO must be set for a student achievement goal, such as a target group of student attaining a particular PARCC score. An example of this would be for 65 percent of 10th grade students to score a four or five on their English 10 PARCC score. The second SLO is to be set on a Climate and Culture goal. An example of this would be to have a 10 percent school-wide reduction of suspensions of African American students, or an increase of 4 percent in the Average Daily Attendance rate for students receiving Free and Reduced Meals (FARMs). Each principal must work with the deputy superintendent to identify their SLO and agree upon the quantitative measures to demonstrate attainment. The deputy superintendent will work with the principals throughout the year once have been agreed upon. There is a timeline for review, with an interim and final meeting outlined with conference topics suggested (SMCPS, 2015). All of this is centered on data gathering, form completion, and compliance with negotiated timelines (SMCPS, 2015). There are only five pages in the manual devoted to SLOs and they constitute 50% of a principal's evaluation. There is no guidance from SMCPS or MSDE as to how an evaluator is to provide support or professional development to the principals as they are expected to improve the performance of their teachers and students. It is in this absence that recent resurgent research into social cognitive theory—specifically, self-efficacy—is being applied to the role of the principal and may well assist those charged with supporting and evaluating school leaders (Tschannen, 2005).

Social Cognitive Theory

When considering the work of principals and how best to evaluate their performance, the personal, behavioral, and environmental influences exerted must be considered (Tschannen, 2005). Central to social cognitive theory is the belief that by selecting environments carefully, we can influence what we become (Nabavi, 2012). "Our choices are influenced by our beliefs as well as our capabilities (and) Bandura proposes only a single internal principle comprised of three interacting elements. This principle is termed triadic reciprocality (p. 14). Bandura defined human behavior as "a triadic, dynamic, and reciprocal interaction of personal factors, behavior, and the environment" (p. 14). Bandura (1997) states that, "reciprocity does not mean that the three sets of interacting determinants are equal strength. Their relative influence will vary for different activities and under different circumstances" (p. 6). Figure 1 is a visual representation of this.

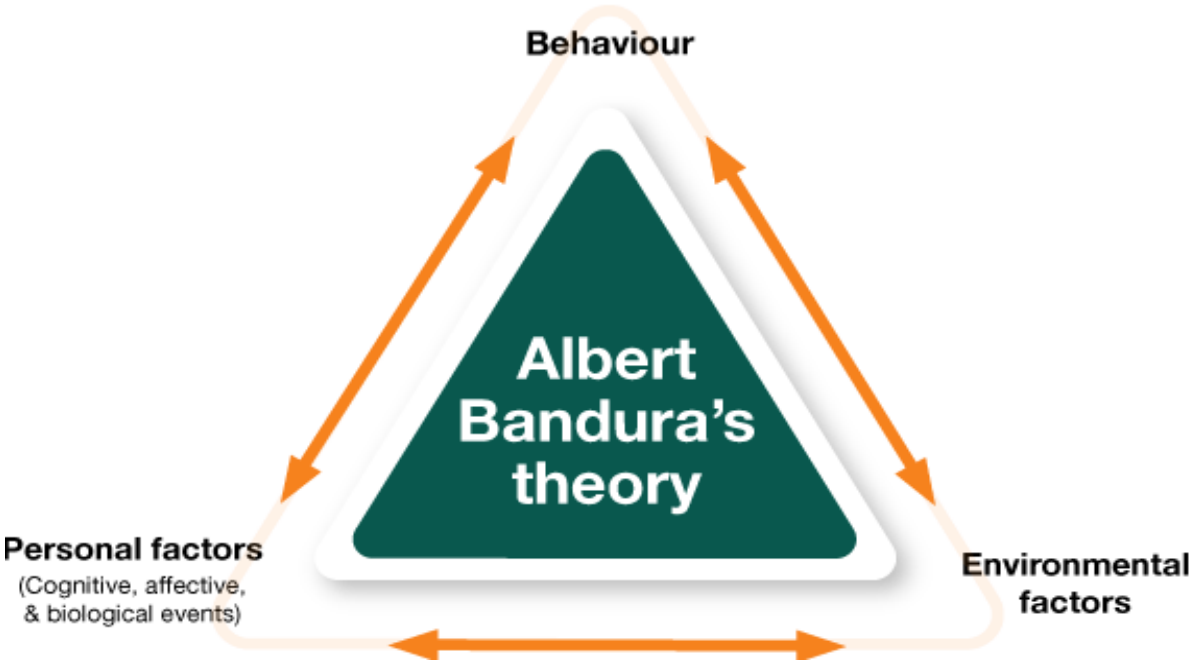

Figure 1. Triadic Reciprocality, Nabavi (2012)

From social cognitive theory came deeper investigation into self-efficacy. Bandura (1997) explored this, stating, "the conceptual and empirical linkages of other determinants to perceived self efficacy deepen understanding of how people guide and shape their own destinies" (p.10). He further states, "the value of a theory is ultimately judged by the power of the methods it yields to effect changes. Self-efficacy theory provides explicit guidelines on how to enable people to exercise some influence over how they live their lives" (p. 10). When evaluators are asked to support principals through the SLO process, it is important to understand the driving forces behind a principal's belief that they can make a meaningful impact. A principal's sense of efficacy is their judgment of their capabilities to organize and execute a particular course of action in order to produce desired outcomes in the school he or she leads (Bandura, 1997).

Self-Efficacy Beliefs in Principals

Principals must have the confidence in the decisions they make as the instructional leader of the school (Tschannen, 2005). Bandura (2000) explained that, "when faced with obstacles, setbacks, and failures, those who doubt their capabilities slacken their efforts, give up, or settle for mediocre solutions. Those who have a strong belief in their capabilities redouble their efforts to master the challenge" (p. 120). This is an essential element of effective SLO goal setting: a principal's belief in their abilities. Combining social cognitive theory and self-efficacy, McCormick (2001) states, "Successful leadership involves using social influence processes to organize, direct, and motivate the actions of others. It requires persistent task-directed effort, effective task strategies, and the artful application of various conceptual, technical, and interpersonal

skills" (p. 28).

Determinants of Self-Efficacy

Bandura (1997) posited that there are four sources of self-efficacy beliefs: "enactive mastery experiences that serve as indicators of capability; vicarious experiences that alter efficacy beliefs through transmission of competencies and comparison with the attainment of others, verbal persuasion and allied types of social influence that one possesses certain capabilities; and physiological and affective states from which people partly judge their capableness, strength, and vulnerability to dysfunction" (p. 79). When one is considering building an evaluation tool for principals these determinants should be considered. How the SLO process incorporates and emphasizes elements that foster and support self-efficacy needs to be fully understood.

Bandura (1997) stated that mastery experiences were "the most influential source of efficacy information because they provide the most authentic evidence whether one can muster whatever it takes to succeed" (p. 80). From his study, he concluded that "the more beliefs of personal efficacy were raised, the better they performed" (p. 81). This has a great deal of relevance for any evaluation system predicated on the attainment of an identified goal because, as he states, "perceived self-efficacy is often a better predictor under variable conditions then past performance, because efficacy judgment encompasses more information than just the executed action" (p. 81).

Beyond the enactive mastery of a task, Bandura identifies vicarious experiences and learning as determinants of self-efficacy beliefs because "people do not rely on inactive experiences as the sole source of information about their capabilities. Efficacy appraisals are partly influenced by vicarious experiences mediated through model

attainments" (p. 86). This has relevance to the procedures employed in the SLO process and the extent to which principals can share their learning with their peers.

Bandura identifies verbal persuasion as another determinant of self-efficacy. Bandura (1997) states "social persuasion serves as a further means of strengthening people's beliefs that they possess the capabilities to achieve what they seek" (p.101). In this way, the coaching and interaction one has with an evaluator can enhance one's performance. Bandura (1997) expressed "it is easier to sustain a sense of efficacy, especially when struggling with difficulties, if significant others expressed faith in one's capabilities" (p.101).

The final determinant of self-efficacy beliefs identified by Bandura is physiological and affective states. According to Bandura (1997), "in judging their capabilities, people rely partly on somatic information conveyed by physiological and emotional states" (p. 106). When one considers the stressors affecting the work of principals, understanding how each handles these factors is important. "The perception of the environment as uncontrollable typically has a debilitating effect on individual goal setting and problem solving" (Tschannen-Moran, 2005, p. 5). The evaluation of principals must consider the elements over which they have control and those over which they do not.

The Need for the Study

Principal evaluation has been evolving over the past several decades and is now focused on quantifying the impact of the principal on student achievement and school culture (Slotnik, Bugler, & Liang, 2015). While "some empirical studies have been conducted on principal self-efficacy there has been no common agreement about how

these constructs should be conceptualized and measured" (p.296, Federici, 2012).

One prominent tool being used is the Principal Self-Efficacy Survey (PSES) instrument developed by Tschannen-Moran and Gareis (2004). The instrument "was developed as an adaptation of the TSES measure presented by Tschannen-Moran and Woolfolk Hoy" (2001). Initially, fifty items were generated to tap various aspects of principals' work. These items, based largely on the professional standards articulated by the Interstate School Leaders Licensure Consortium (ISLLC), were then submitted to a panel of experts, which included three professors of educational leadership and one practicing superintendent, for review and refinement. Next, the instrument was field tested with ten former principals to check for the clarity of directions, appropriateness of the items and response scale, and any other observations or feedback they were willing to share (Tschannen-Moran & Gareis, 2004, p. 579). "Using principal axis factor analysis, the 50 original items of the PSES were reduced to a scale with 18 items" (Tschannen-Moran & Gareis, 2004, p. 580).

For this study, the PSES will be used in a new context, which adds to a body of knowledge about this scale. Linking PSES survey results with a series of questions about the SMCPS SLO process, aligned to Bandura's determinants of self-efficacy, may inform how social cognitive theory impacts principal motivation related to the implementation of the SLO. The additional research questions are needed to help evaluators further refine the SLO process to best support principals.

Section II: Investigation—Generating and Defending an Original Solution

Purpose of the Study

The purpose of this directed study was to contribute to the research about principal self-efficacy beliefs and gather principals' reactions to their evaluation experiences relative to the SLO process thus informing the practice in the school district. Participants completed the PSES followed by a series of questions eliciting reactions to the SLO process from the principals who have gone through an evaluation cycle. The specific questions asked of principals, relative to their SLO process, were aligned with Bandura's four determinants of self-efficacy: mastery learning, vicarious learning, verbal persuasion, and emotional arousal. This research is important to leaders in Maryland, as much has been invested in the SLO process and if this tool positively impacts self-efficacy beliefs of principals, this may have a cascading influence on the self-efficacy beliefs of teachers and students, which has been shown to impact achievement. By learning more about principal self-efficacy beliefs and factors which positively or negatively influence such beliefs, recommendations can be made for how to best tailor the SLO process to increase principal self-efficacy beliefs.

Research Questions

Albert Bandura developed the concept of self-efficacy as one component of social cognitive theory (Bandura, 1977). Positive self-efficacy beliefs lead people to feel they can produce behaviors that will lead to future success (Dimmock & Hattie, 1996). Additionally, goal setting is strongly influenced by levels of self-efficacy beliefs, higher levels of efficacy beliefs lead to the establishment of more challenging goals and stronger

level of commitment and perseverance (Bandura, 1993, & Zulkosky, 2009). With this accepted, it is worth exploring if the current principal evaluation system, predicated on SLOs, has a relationship to the self-efficacy beliefs of participants

Bandura identified four main sources of self-efficacy beliefs: mastery learning, vicarious learning, verbal persuasion, and emotional arousal (Bandura, 1997, & Bandura, 1993). Bandura posits that self-efficacy increases by successfully completing tasks, seeing someone else complete tasks, receiving positive feedback, or believing that a situation is relevant and meaningful to them (Bandura, 1997). Using this as an organizational framework, the following research questions were posed:

1. Does the attainment of a principal's SLO for Student Achievement have a significant relationship with the principal's level of self-efficacy (PSES mean score)?
2. Does the attainment of a principal's SLO for Culture and Climate have a significant relationship with the principal's level of self-efficacy (PSES mean score)?
3. Does the attainment of a principal's SLO for Student Achievement have a significant relationship with the principal's mean score for questions regarding enactive mastery learning, vicarious learning, verbal persuasion, and physiological arousal?
4. Does the attainment of a principal's SLO for Culture and Climate have a significant relationship with the principal's mean score for questions regarding enactive mastery learning, vicarious learning, verbal persuasion, and physiological arousal?

Hypotheses

Two hypotheses identified for testing by this study were:

1. Principals attaining their Student Learning Outcome for Student Achievement will have a higher mean PSES score overall and across all three sublevel constructs and a higher mean score on questions aligned to Bandura's determinants of self-efficacy: mastery level learning, vicarious learning, verbal persuasion, and physiological arousal.
2. Principals attaining their Student Learning Outcome for Climate and Culture will have a higher mean PSES score overall and across all three sublevel constructs and a higher mean score on questions aligned to Bandura's determinants of self-efficacy: mastery level learning, vicarious learning, verbal persuasion, and physiological arousal.

Methods and Procedures

Participants

The school system has eighteen elementary schools, four middle schools, three high schools, two academies, and a technical center. The study was sent to all twenty-eight educational leaders as they have all gone through the evaluation process for the 2016–2017 school year and had SLOs recorded as 50% of their overall evaluation.

Instrument

The research design of this directed study is quantitative in nature and was conducted through a questionnaire or survey. The participants in this study were asked to complete the PSES. The PSES survey consisted of eighteen questions asking respondents

to respond according to a nine-point Likert scale from "not at all" to "a great deal." The research instrument utilized in this study was that of questionnaires that were administered to the principals in this study. Table 1 illustrates the possible answers and range of responses in the Likert-type survey used for this study.

Table 1. Response scale.

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

Following the eighteen questions from the PSES, respondents were asked twelve additional questions regarding the SLO process with the responses framed using the same scale. The proposed instrument is located in Appendix A. The last two questions asked of respondents were whether or not they attained their proposed SLOs—one for student achievement and one for climate and culture.

Survey Structure

The survey is structured into three sections, with the Section I comprised of the 18 questions from the Principal Self-Efficacy survey, which can be further broken into three sub-level constructs of instructional, moral, and managerial leadership. Section II is comprised of 12 questions that can also be broken into four sub-level constructs aligned to the four determinants of Albert Bandura's determinants of self-efficacy, mastery learning, vicarious learning, verbal persuasion, and physiological arousal and the affective state. The final two questions about attainment of the SLOs for student achievement and culture and climate will act as the sorting group. Figure 3 represents this structure.

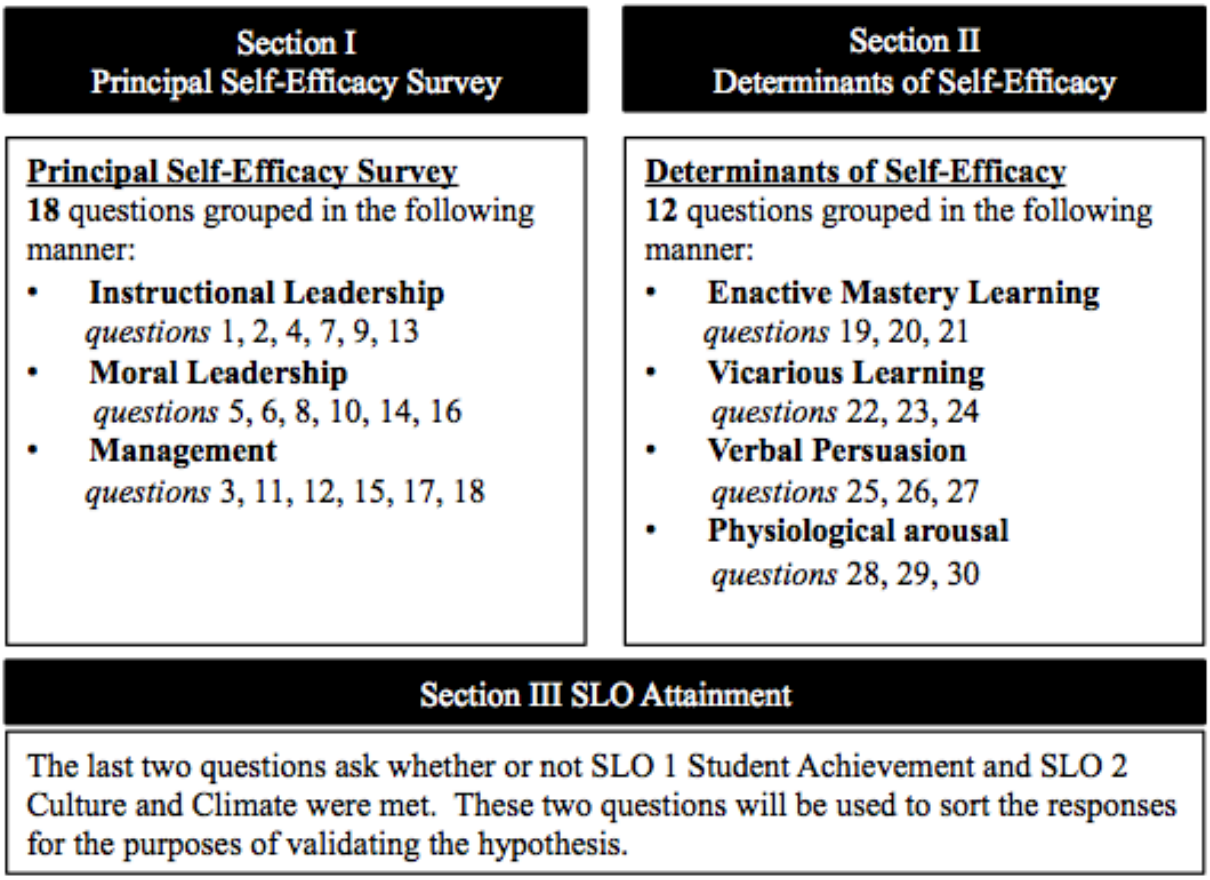

Figure 3. Structure of the complete survey

Pilot Testing the Instrument

Prior to beginning the study, the survey was pilot tested. The survey was distributed for feedback to four individuals who were former principals and have experience with the SLO process. After administering the pilot survey, each question was reviewed with the participants. Adjustments were made to the final two questions regarding their attainment of their SLOs and simplified to a simple yes or no response.

Data Collection

The surveys were distributed to all current principals in the school system using the Qualtrics platform. An email was sent to principals by the system's Administrative

Accountability Officer, explaining the purpose of the study and asking for their participation by completing a survey that was accessible through a link provided in the body of the email. The link took respondents to a "Consent to Participate" page that required them to provide their statement of consent by clicking "Yes, I voluntarily agree to participate" or "No, I do not agree to participate." They were asked to respond within the next three weeks. The email also informed them that there would be an incentive of a $5 gift card to Starbucks after completing the survey. After the first week, a reminder email was sent to all principals each week the survey was open. All responses were completely anonymous and no data were gathered regarding the respondent's gender, assignment, tenure, or educational level.

Analysis of Data

Data obtained from the online surveys were analyzed by using the Qualtrics software platform to determine the minimum, maximum, mean, standard deviation, variance, and count for each question on the survey. The data were also cross tabulated according to respondents who attained their SLOs compared to those who did not.

Limitations

As with any research study, limitations are recognized. The following limitations were identified for this study.

1. Limitations with Likert scales—ordinal verse interval. "Maurer and Pierce (1998) investigated the effectiveness of a Likert scale measure of self-efficacy for academic performance. They suggested that the Likert scale can be considered a measure of both magnitude and confidence, and they concluded,

based on reliability, predictive validity, and factor analysis data, that a Likert scale measure of self-efficacy is an acceptable alternative to the traditional measure. In that study, the two scales provided similar results. Maurer and Pierce suggested that the Likert scale may provide a simpler method for assessing self-efficacy (Maurer, 2000).

2. Small sample size. Small sample sizes may influence the validity of the study. There are only twenty-eight principals in SMCPS and, although there was a very high rate of response (twenty-six out of twenty-eight), that is still a fairly small sample size.

3. Self-reported data. As with any survey that relies on the respondents to self-report, the data may be biased to the perception of the individual.

4. Sensitivity of the instrument. Although the PSES has been used in multiple research studies, the following questions regarding the SLO process in SMCPS had not. The four hypotheses linking participant's PSES responses to questions about the SLO process had been done before, so there is no proof of concept.

Although the questionnaire was anonymous and confidential, subjects may have had concerns that their responses could be linked to them. The researcher adhered to the guidelines for the University of Maryland Institutional Review Board (IRB). The following procedures were used to ensure that the identities of all respondents remained confidential.

- To prevent an infringement on confidentiality, the individual data collected through the online survey will not be made public, only averages of responses.
- Each respondent was provided access to the survey, and the researcher was the only person with access to the questionnaire data.
- All participants received an email describing the study and detailing their confidentiality in participation in the study.
- All participants had to agree with an informed "consent to participate" electronically before beginning the survey.
- To maintain confidentiality, the researcher did not gather any identifying information from respondents.
- The researcher retained the data from the surveys electronically on a personal computer, and the researcher has sole access to the information contained on said computer.
- All data will be erased after five years.

Sampling

The sampling method utilized in this study is the purposive sampling method. The survey was piloted informally, using four central office administrators. The survey was refined based on their feedback. Once the instrument was finalized and the time to complete it assessed, it was sent to principals.

Summary

The purpose of this directed study was to contribute to the research about principal self-efficacy beliefs and gather principals' reactions to their evaluation

experiences relative to the SLO process thus informing the practice in the school district. Participants completed the PSES followed by a series of questions eliciting reactions to the SLO process from the principals who have gone through an evaluation cycle. As discussed in Section I, a great deal of effort has gone into principal and teacher evaluation in Maryland over the past decade. Much work has been compelled by legislation and federal mandate, but all of it has been driven by the need to improve public education and increase student achievement. School principals are at the center of this work and a more complete and nuanced understanding of their role and how best to support them is essential. As discussed, the NAESP and the NASSP have stated that a principal evaluation system must be:

1. Created by and for principals;
2. Part of a comprehensive system of support and professional development;
3. Flexible enough to accommodate differences in principals' experiences;
4. Relevant to the improvement of principals' dynamic work;
5. Based on accurate, valid, and reliable information gathered through multiple measures;
6. Fair in placing a priority on outcomes that principals can control; and
7. Useful for informing principals' learning and progress (Clifford, M., & Ross, 2013).

Bandura's determinants of self-efficacy can be found in each of the identified elements. Mastery learning drives the sequential outcomes of a comprehensive system. Vicarious learning can be seen in the expectation that the system be built by principals and predicated on professional development. Verbal persuasion is built into the

flexibility, fairness, and systems of support. The final determinant, physiological arousal, is found in the system's usefulness, relevance, and the priorities placed on outcomes the principals can control. Simply put, to exclude examinations of a principal's self-efficacy beliefs is to overlook a primary driver of how principals approach their work and move others.

As set out in Section II, the PSES was used in a new context, which adds to a body of knowledge about this scale. The results from the PSES provided a baseline for self-efficacy beliefs of the system's principals and provided comparative data when considering the second set of questions regarding the SLO process in SMCPS. Linking PSES survey results with a series of questions about the SMCPS SLO process, aligned to Bandura's four main sources of self-efficacy beliefs: mastery learning, vicarious learning, verbal persuasion, and emotional arousal, provided a deeper insight into the SLO process and may inform evaluators how to better support principals and foster stronger beliefs of self-efficacy in principals.

Section III. Results and Conclusions

The purpose of this study was to examine how the inclusion of SLOs has impacted educational leaders in a mid-size, Mid-Atlantic state school system, specifically through the lens of a principal's self-efficacy beliefs. Using a Likert scale survey, broken into three sections, this study served to answer four essential questions:

1. Does the attainment of a principal's SLO for Student Achievement have a significant relationship with the principal's level of self-efficacy, in total and broken out by the three sublevel constructs: instructional leadership, moral leadership, and managerial leadership, as measured by the PSES?
2. Does the attainment of a principal's SLO for Culture and Climate have a significant relationship with the principal's level of self-efficacy, in total and broken out by the three sublevel constructs: instructional leadership, moral leadership, and managerial leadership, as measured by the PSES?
3. Does the attainment of a principal's SLO for Student Achievement have a significant relationship with the principal's mean score for questions regarding enactive mastery learning, vicarious learning, verbal persuasion, and physiological arousal?
4. Does the attainment of a principal's SLO for Culture and Climate have a significant relationship with the principal's mean score for questions regarding enactive mastery learning, vicarious learning, verbal persuasion, and physiological arousal?

A total of 26 out of 28 possible participants responded to the survey, a response rate of 92.85%. Of the 26 responding, all but three answered all 34 questions. One

participant did not respond to question 7, one did not respond to question 9, and one did not respond to question 16. There is no way to determine if the non-responses were intentional or inadvertent. The survey window was open for three weeks, through late August and the first week of September of 2017, with participants being able to complete the survey on a computer, tablet, or mobile device.

To answer the four questions posed in this research, the researcher ran the responses through the Qualtrics database to determine the minimum, maximum, mean, standard deviation, variance, and count for each of the Likert questions in Section I and Section II of the survey. These data were then exported to an Excel spreadsheet for further analysis. Once in an Excel spreadsheet, the responses were organized according to the PSES sublevel constructs and the four determinants of self-efficacy determinants as defined by Albert Bandura. Once the data were organized in this fashion, responses were averaged into mean scores for analysis. For the purposes of hypothesis testing, the level of statistical significance of the study was .05. All data analysis and graph construction for the study took place in Stata / SE 14.2 software.

Survey Section I. Principal Self-Efficacy Survey

The 18 items of the PSES, developed by Tschannen-Moran and Gareis (2004), were the first questions asked of the respondents and they were asked to indicate their opinion about each of the questions by marking one of the nine responses along a scale ranging from "None at All" (1) to "A Great Deal" (9), with "Some Degree" (5) representing the middle between these low and high values. Participants were directed to consider the combination of their current ability, resources, and opportunity to do each of the following in their position last year. Each question asked began with, "In your current

role as principal, to what extent can you …" followed by the specific question. For example, the first question was "In your current role as principal, to what extent can you facilitate student learning in your school?"

The questions on the PSES are grouped into three sublevel constructs: instructional leadership, managerial leadership, or moral leadership. Each area is examined by six survey items.

The questions designed to explore **instructional leadership** are as follows:

In your current role as a principal, to what extent can you…

- facilitate student learning in your school?
- generate enthusiasm for a shared vision for the school?
- manage change in your school?
- create a positive learning environment in your school?
- raise student achievement on standardized tests?
- motivate teachers?

The questions related to **managerial leadership** are the following:

In your current role as a principal, to what extent can you…

- handle the time demands of the job?
- maintain control of your own daily schedule?
- shape the operational policies and procedures that are necessary to manage your school?
- handle the paperwork required of the job?
- cope with the stress of the job?
- prioritize among competing demands of the job?

The questions related to **moral leadership** are the following:

In your current role as a principal, to what extent can you…

- promote school spirit among a large majority of the student population?
- promote a positive image of your school with the media?
- promote the prevailing values of the community in your school?
- handle effectively the discipline of students in your school?
- promote acceptable behavior among students?
- promote ethical behavior among school personnel?

To answer Research Questions and 1 and 2, the means and standard deviations were calculated for the total self-efficacy scale and for the subscales that measure efficacy in instructional leadership, managerial leadership, and moral leadership.

The composite mean score for the 26 principals completing the PSES was 134.08 out of a possible 162 points. When examining the sublevel constructs, out of a possible 54 points, the highest mean score for the subscales was in moral leadership (M = 47.08), followed by instructional leadership (M = 44.85) and managerial leadership (M = 42.15). These data are depicted in Table 2.

Table 2. *Means, Standard Deviation, and Ranges for Total PSES and Subscales*

Scale	Mean	Standard Deviation	Range
Total Scale	134.08	11.08	128–158
Instructional Leadership	42.15	4.01	36–54
Managerial Leadership	42.15	6.25	35–51

Moral Leadership	47.08	3.96	41–54

Research Question 1 Results: Overall PSES Score and SLO 1

The purpose of this section is to present the results of the study. The results have been divided into three parts. First, the research questions and hypotheses of the study were restated. Second, the answers to the research questions were presented. Third, the findings of the study were summarized by means of a separate evaluation of each of the null hypotheses. The first research question explored whether or not the overall PSES score and the attainment of SLO 1 had any significance.

RQ1: Is there a significant difference between mean **overall self-efficacy scores** for (a) principals who met the SLO 1 standard and (b) principals who did not meet the SLO 1 standard?

HI_0: Mean overall self-efficacy scores for principals who met the SLO 1 standard = Mean overall self-efficacy scores for principals who did not meet the SLO 1 standard.

H$I I_A$: Mean overall self-efficacy scores for principals who met the SLO 1 standard ≠ Mean overall self-efficacy scores for principals who did not meet the SLO 1 standard.

For the purposes of hypothesis testing, the level of statistical significance of the study was .05. All data analysis and graph construction for the study took place in Stata / SE 14.2 software. Scores on overall self-efficacy were contrasted based on whether principals met or failed to meet the SLO 1 standard. There were eighteen principals who reported meeting the SLO 1 standard and eight principals who reported not meeting the

SLO 1 standard. Because the sorting group (meeting or not meeting the SLO 1 standard) was dichotomous and independent and because the dependent variable of score was continuous, both an independent samples *t*-test and its non-parametric equivalent, the Mann-Whitney U test, were applied.

The independent samples *t*-test indicated that the difference, -0.15, between overall self-efficacy scores for principals passing SLO 1 (M = 75.28, SD = 14.27, SEM = 3.36, n = 18) and overall self-efficacy scores for principals not passing SLO 1 (M = 75.13, SD = 13.81, SEM = 4.88, n = 8) was not statistically significant, $t(24) = -0.03$, p = 0.98. The 95% confidence interval of the difference between these two groups, -12.55 to 12.25, included the value of 0, providing another means to infer that there was no significant overall self-efficacy difference between principals passing SLO 1 and principals not passing SLO 1. These results were triangulated by the Mann-Whitney U test, U = 70.50, z = 0.06, p = .95. Therefore, by both parametric and non-parametric means, it was established that there was no significant difference in the overall self-efficacy scores of principals who passed SLO 1 and principals who did not pass SLO 1.

The box plot below in Figure 3 contains the comparison of mean overall self-efficacy scores for principals who passed SLO 1 and principals who did not pass SLO 1. The box plot clearly indicates the existence of a substantial overlap in the distributions of overall self-efficacy scores for principals who passed SLO 1 and principals who did not pass SLO 1.

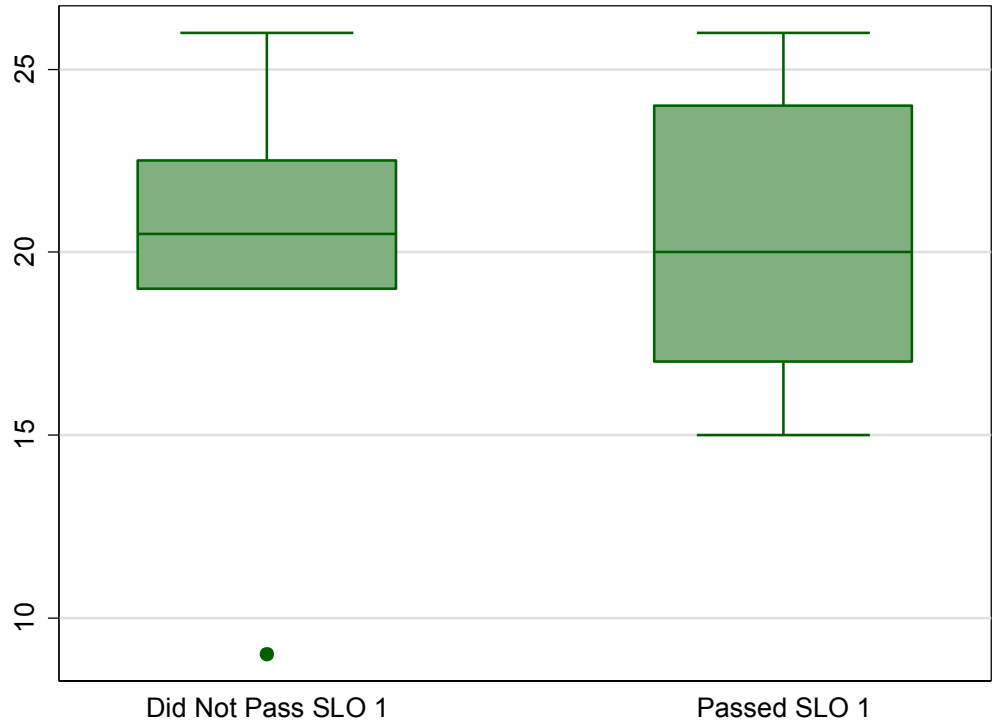

Figure 3. Box plot, comparison of overall self-efficacy by SLO 1 status.
Note: Original graphic generated in Stata / SE 14.2.

The null hypothesis for overall self-efficacy differences between principals who passed SLO 1 and principals who did not pass SLO 1 was that mean overall self-efficacy would be identical between these two groups. On the basis of both the independent samples *t*-test and the Mann-Whitney U test, whose results have been reported above for overall self-efficacy as related to SLO 1 passing status, the null hypothesis could not be rejected. There was insufficient evidence, at an Alpha of .05, for the existence of a significant difference in SLO 1 scores between principals who passed SLO 1 and principals who did not pass SLO 1.

After overall analysis did not yield significant statistical difference in the PSES score and the variable of SLO 1 attainment, further research was done by running the same set of statistics for each of the sublevel constructs of the PSES. The research

questions for this are stated as such:

RQ1.1: Is there a significant difference between mean **instructional leadership** scores for (a) principals who met the SLO 1 standard and (b) principals who did not meet the SLO 1 standard?

- H*1.1* $_0$: Mean effective mastery scores for principals who met the SLO 1 standard = Mean effective mastery scores for principals who did not meet the SLO 1 standard.
- H*1.1* $_A$: Mean effective mastery scores for principals who met the SLO 1 standard ≠ Mean effective mastery scores for principals who did not meet the SLO 1 standard.

RQ1.2: Is there a significant difference between mean **moral leadership** scores for (a) principals who met the SLO 1 standard and (b) principals who did not meet the SLO 1 standard?

- H*1.2*$_0$: Mean vicarious learning scores for principals who met the SLO 1 standard = Mean vicarious learning scores for principals who did not meet the SLO 1 standard.
- H*1.2*$_A$: Mean vicarious learning scores for principals who met the SLO 1 standard ≠ Mean vicarious learning scores for principals who did not meet the SLO 1 standard.

RQ1.3: Is there a significant difference between mean **managerial leadership** scores for (a) principals who met the SLO 1 standard and (b) principals who did not meet the SLO 1 standard?

- H1.3_0: Mean verbal persuasion scores for principals who met the SLO 1 standard = Mean verbal persuasion scores for principals who did not meet the SLO 1 standard.
- H1.3_A: Mean verbal persuasion scores for principals who met the SLO 1 standard ≠ Mean verbal persuasion scores for principals who did not meet the SLO 1 standard.

Research Question 1.1 Results: Instructional Leadership and SLO 1

The difference in **instructional leadership** between those who met SLO 1 and those who did not meet SLO 1 was also measured through an independent samples *t*-test. The mean instructional leadership for those who met SLO 1 (M = 45.50, SD = 4.43, n = 18) was compared to the mean instructional leadership for those who did not meet SLO 1 (M = 43.38 SD = 2.44, n =8). The calculated difference of -2.13 (95% confidence interval = -5.60 to 1.35) was found to not be statistically significant, $t(24) = 1.22$, $p = 0.23$.

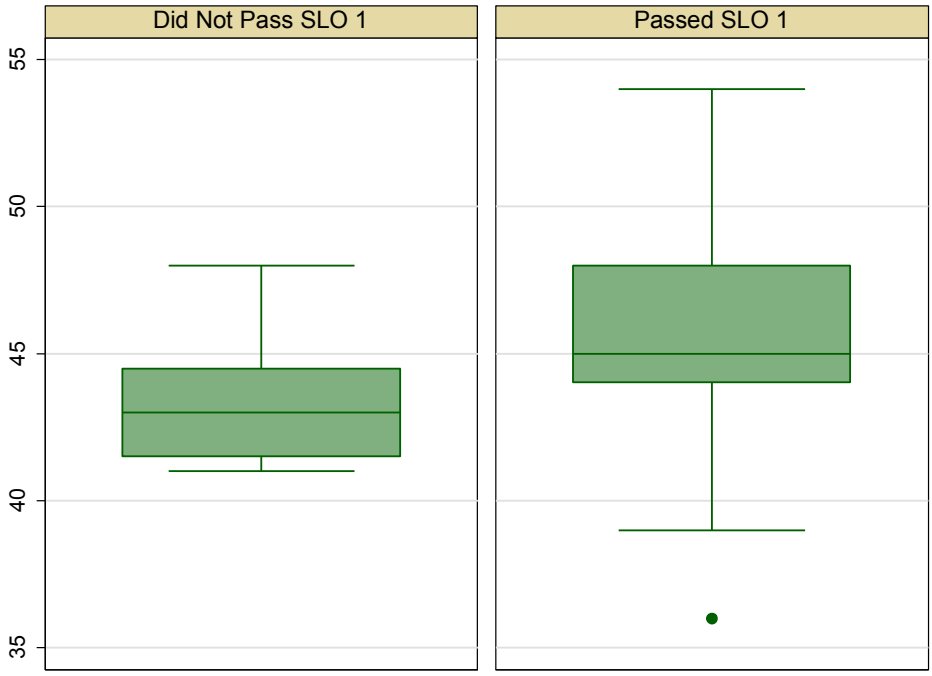

Figure 4. Box plot, Instructional leadership by SLO 1 status. *Note*: Original graphic generated in Stata / SE 14.2.

The Mann-Whitney U test for the difference in instructional leadership by mean SLO 1 pass status was also not significant, $MWU = 43$, $z = 1.58$, $p = .11$. Therefore, by both parametric and non-parametric techniques, there was found to be no statistically significant effect of SLO 1 passing status on differences in instructional leadership.

Research Question 1.2 Results: Moral Leadership and SLO 1

Next, the mean **moral leadership** for those who met SLO 1 ($M = 47.61$, $SD = 4.37$, $n = 18$) was compared to the mean moral leadership for those who did not meet SLO 1 ($M = 45.88$ $SD = 2.70$, $n = 8$). The calculated difference of -1.74 (95% confidence interval = -5.20 to 1.73) was found to not be statistically significant, $t(24) = 0.95$, $p = .35$. The Mann-Whitney U test for the difference in moral leadership by mean SLO 1 pass status was also not significant, $MWU = 52.5$, $z = 1.06$, $p = .29$. Therefore, by both

parametric and non-parametric techniques, there was found to be no statistically significant effect of SLO 1 passing status on differences in moral leadership.

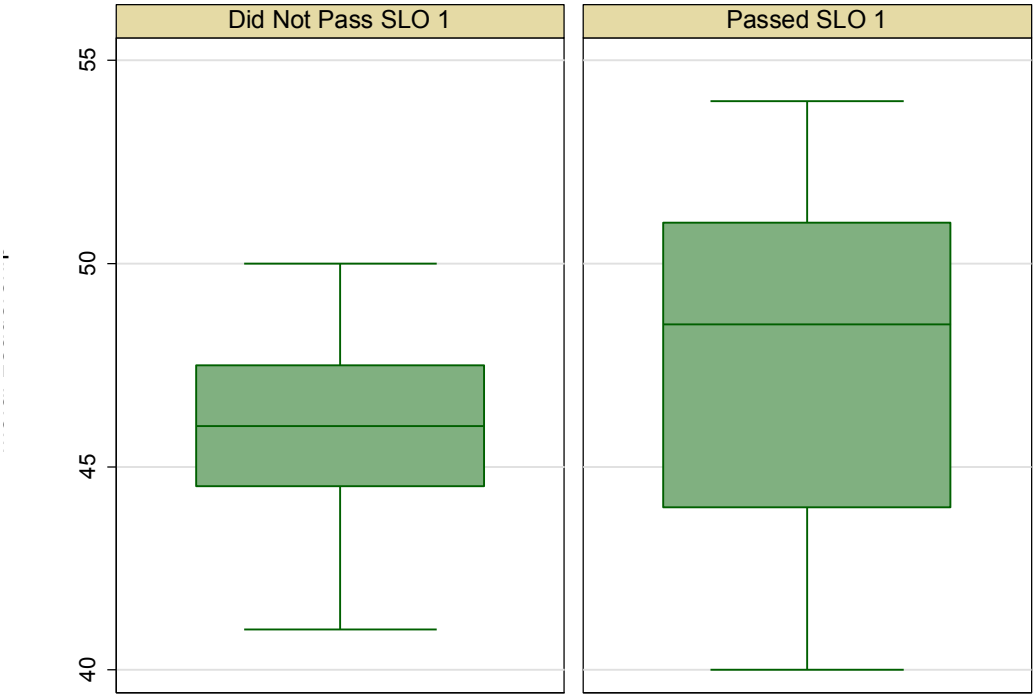

Figure 5. Box plot, Moral leadership by SLO 1 status. *Note*: Original graphic generated in Stata / SE 14.2

Research Question 1.3 Results: Managerial Leadership and SLO 1

Finally, the mean management score for those who met SLO 1 (M = 42.28, SD = 6.84,

n = 18) was compared to the mean moral leadership for those who did not meet SLO 1 (M = 41.88 SD = 5.03, n =8). The calculated difference of -0.40 (95% confidence interval = -5.99 to 5.18) was found to not be statistically significant, $t(24) = 0.12$, $p = .90$. The Mann-Whitney U test for the difference management by mean SLO 1 pass status was also not significant, $MWU = 60$, $z = 0.64$, $p = .52$. Therefore, by both parametric and non-parametric techniques, there was found to be no statistically significant effect of SLO 1

passing status on differences in management.

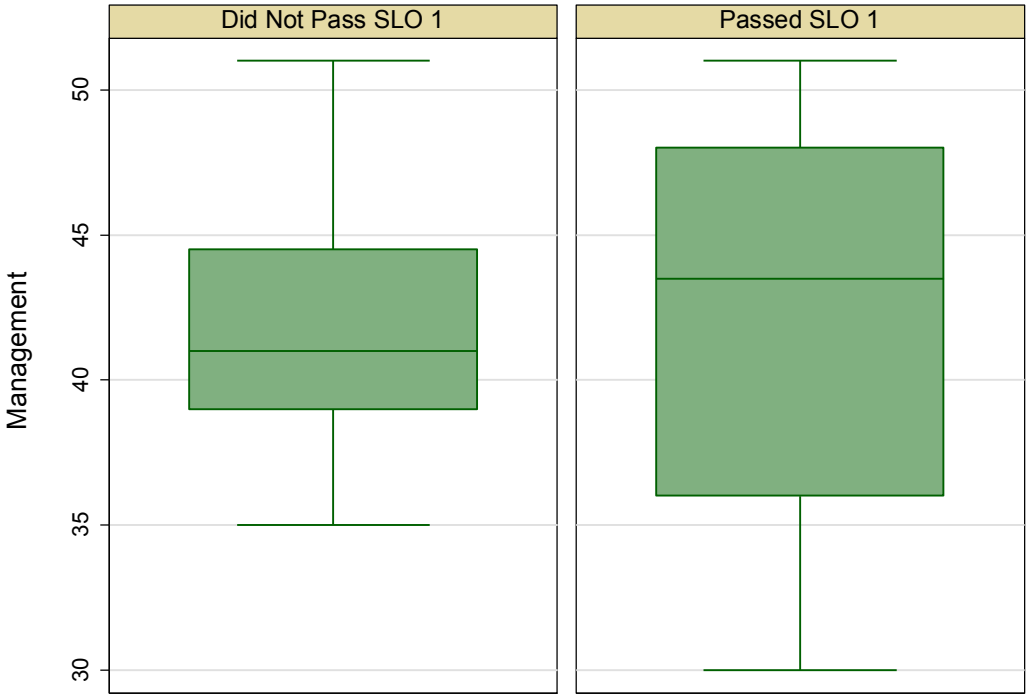

Figure 6. Box plot, Managerial leadership by SLO 1 status. *Note*: Original graphic generated in Stata / SE 14.2

Research Question 2 Results: Overall PSES Score and SLO 2

Scores on overall self-efficacy were contrasted based on whether principals met or failed to meet the SLO 2 standard. There were 20 principals who reported meeting the SLO 2 standard and six principals who reported not meeting the SLO 2 standard. The research questions were stated as such:

RQ2: Is there a significant difference between mean **overall self-efficacy scores** for (a) principals who met the SLO 2 standard and (b) principals who did not meet the SLO 2 standard?

$H2_0$: Mean overall self-efficacy scores for principals who met the SLO 2 standard = Mean overall self-efficacy scores for principals who did not meet the SLO 2

standard.

$H2_A$: Mean overall self-efficacy scores for principals who met the SLO 2 standard ≠ Mean overall self-efficacy scores for principals who did not meet the SLO 2 standard.

For the purposes of hypothesis testing, the level of statistical significance of the study was .05. All data analysis and graph construction for the study took place in Stata / SE 14.2 software. Scores on overall self-efficacy were contrasted based on whether principals met or failed to meet the SLO 2 standard. Because the sorting group (meeting or not meeting the SLO 2 standard) was dichotomous and independent, and because the dependent variable of score was continuous, both an independent samples *t*-test and its non-parametric equivalent, the Mann-Whitney U test, were applied.

The independent samples *t*-test indicated that the difference, 3.38, between overall self-efficacy scores for principals passing SLO 2 ($M = 74.45$, $SD = 14.05$, $SEM = 3.14$, $n = 20$) and overall self-efficacy scores for principals not passing SLO 2 ($M = 77.83$, $SD = 14.10$, $SEM = 5.76$, $n = 6$) was not statistically significant, $t(24) = 0.52$, $p = .61$. The 95% confidence interval of the difference between these two groups, -10.13 to 16.89, included the value of 0, providing another means to infer that there was no significant overall self-efficacy difference between principals passing SLO 2 and principals not passing SLO 2. These results were triangulated by the Mann-Whitney U test, $U = 48.50$, $z = -0.67$, $p = .50$. Therefore, by both parametric and non-parametric means, it was established that there was no significant difference in the overall self-efficacy scores of principals who passed SLO 2 and principals who did not pass SLO 2.

The box plot in Figure 7 below contains the comparison of mean overall self-

efficacy scores for principals who passed SLO 2 and principals who did not pass SLO 2. The box plot clearly indicates the existence of a substantial overlap in the distributions of overall self-efficacy scores for principals who passed SLO 2 and principals who did not pass SLO 2.

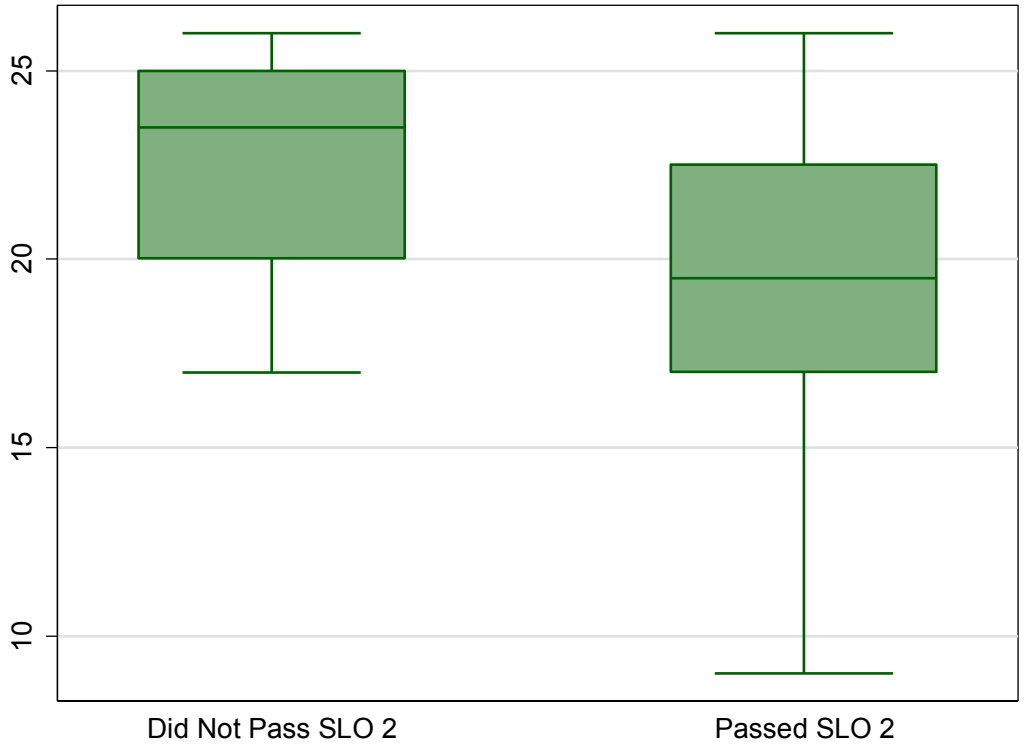

Figure 7. Box plot, comparison of overall self-efficacy by SLO 2 status. *Note*: Original graphic generated in Stata / SE 14.2.

The null hypothesis for overall self-efficacy differences between principals who passed SLO 2 and principals who did not pass SLO 2 was that mean overall self-efficacy would be identical between these two groups. On the basis of both the independent samples *t*-test and the Mann-Whitney U test, whose results have been reported above for overall self-efficacy as related to SLO 2 passing status, the null hypothesis could not be rejected. There was insufficient evidence, at an Alpha of .05, for the existence of a significant difference in SLO 2 scores between principals who passed SLO 2 and

principals who did not pass SLO 2.

Again, overall analysis did not yield significant statistical difference in the PSES score and the variable of SLO 2 attainment, further research was done by running the same set of statistics for each of the sublevel constructs of the PSES. The research questions for this are stated as such:

RQ2.1: Is there a significant difference between mean **instructional leadership** scores for (a) principals who met the SLO 2 standard and (b) principals who did not meet the SLO 2 standard?

- H*2.1*$_0$: Mean effective mastery scores for principals who met the SLO 2 standard = Mean effective mastery scores for principals who did not meet the SLO 2 standard.
- H*2.1*$_A$: Mean effective mastery scores for principals who met the SLO 2 standard ≠ Mean effective mastery scores for principals who did not meet the SLO 2 standard.

RQ2.2: Is there a significant difference between mean **moral leadership** scores for (a) principals who met the SLO 2 standard and (b) principals who did not meet the SLO 2 standard?

- H*2.2*$_0$: Mean vicarious learning scores for principals who met the SLO 1 standard = Mean vicarious learning scores for principals who did not meet the SLO 1 standard.
- H*2.2*$_A$: Mean vicarious learning scores for principals who met the SLO 1 standard ≠ Mean vicarious learning scores for principals who did not meet the SLO 1 standard.

RQ2.3: Is there a significant difference between mean **managerial leadership** scores for (a) principals who met the SLO 2 standard and (b) principals who did not meet the SLO 2 standard?

- H*2.3*$_0$: Mean verbal persuasion scores for principals who met the SLO 2 standard = Mean verbal persuasion scores for principals who did not meet the SLO 2 standard.
- H*2.3*$_A$: Mean verbal persuasion scores for principals who met the SLO 2 standard ≠ Mean verbal persuasion scores for principals who did not meet the SLO 2 standard.

Research Question 2.1 Results: Instructional Leadership and SLO 2

The difference in instructional leadership between those who met SLO 2 and those who did not meet SLO 2 was measured through an independent samples *t*-test. The mean instructional leadership for those who met SLO 2 ($M = 44.95$, $SD = 4.34$, $n = 20$) was compared to the mean instructional leadership for those who did not meet SLO 2 ($M = 44.50$ $SD = 2.88$, $n = 6$). The calculated difference of -0.45 (95% confidence interval = -4.37 to 3.47) was found to not be statistically significant, $t(24) = -0.24$, $p = 0.81$. The Mann-Whitney U test for the difference in instructional leadership by mean SLO 2 pass status was also not significant, $MWU = 55$, $z = 0.27$, $p = 0.79$. Therefore, by both parametric and non-parametric techniques, there was found to be no statistically significant effect of SLO 2 passing status on differences in instructional leadership.

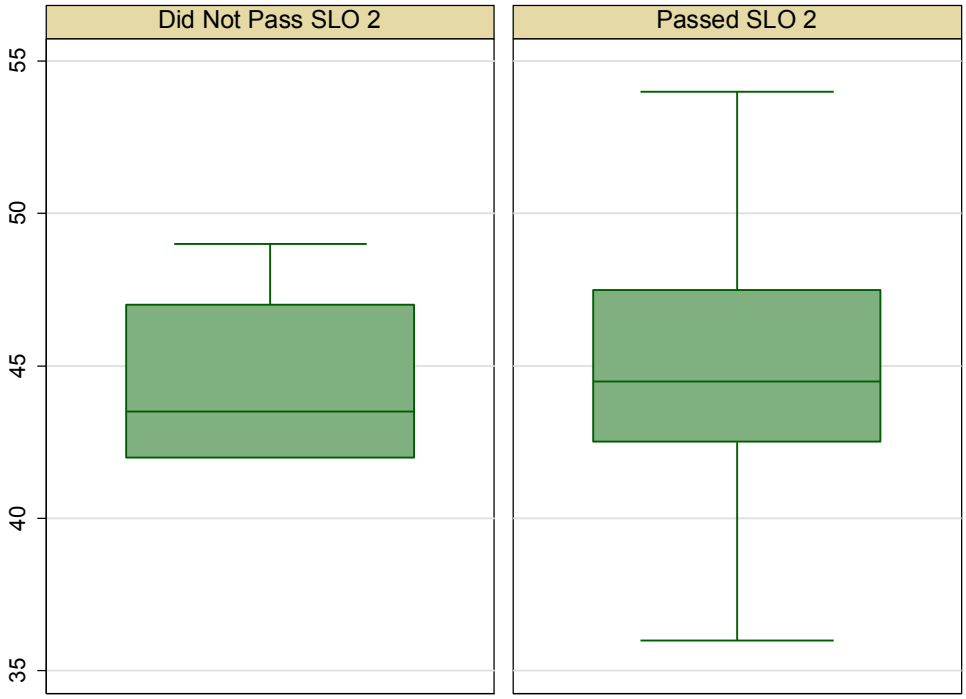

Figure 8. Box plot, Instructional leadership by SLO 2 status. *Note*: Original graphic generated in Stata / SE 14.2

Research Question 2.2 Results: Moral Leadership and SLO 2

The difference in moral leadership between those who met SLO 2 and those who did not meet SLO 2 was measured through an independent samples *t*-test. The mean moral leadership for those who met SLO 2 ($M = 46.30$, $SD = 4.03$, $n = 20$) was compared to the mean moral leadership for those who did not meet SLO 2 ($M = 49.67$ $SD = 2.50$, $n = 6$). The calculated difference of 3.37 (95% confidence interval = -0.25 to 6.98) was found to be statistically significant, such that the mean moral leadership of those who did not pass SLO 2 was higher than the mean moral leadership of those who passed SLO 2, $t(24) = 1.92$, $p = 0.03$. The Mann-Whitney U test for the difference in moral leadership by mean SLO 2 pass status was also significant, $MWU = 29$, $z = -1.86$, $p = 0.03$. Therefore, by both parametric and non-parametric techniques, there was found to be a

statistically significant effect of SLO 2 passing status on differences in moral leadership.

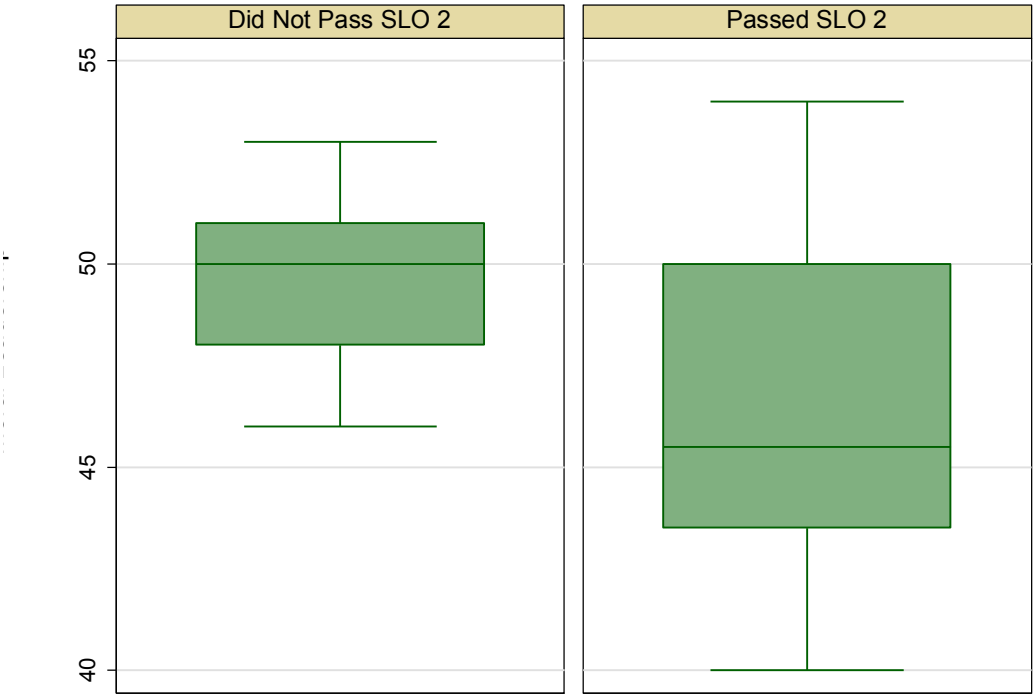

Figure 9. Box plot, Moral leadership by SLO 2 status. *Note*: Original graphic generated in Stata / SE 14.2

Research Question 2.3 Results: Managerial Leadership and SLO 2

The mean management for those who met SLO 2 (M = 42.00, SD = 5.68, n = 20) was compared to the mean management for those who did not meet SLO 2 (M = 42.67, SD = 3.45,

n = 6). The calculated difference of 0.67 (95% confidence interval = -5.44 to 6.78) was found to not be statistically significant, $t(24) = 0.23$, $p = 0.82$. The Mann-Whitney U test for the difference in management by mean SLO 2 pass status was also not significant, $MWU = 51$, $z = -0.52$, $p = 0.30$. Therefore, by both parametric and non-parametric techniques, there was found to be no statistically significant effect of SLO 2 passing status on differences in management.

Figure 10. Box plot, Managerial leadership by SLO 2 status. *Note*: Original graphic generated in Stata / SE 14.2

Summary of Findings for PSES and SLO Attainment

The findings of the first section of the study have been summarized with respect to the two research questions posed and the three sublevel constructs contained in each. The findings have been discussed with respect to existing theories and empirical findings in the final section of the study. The main point requiring discussion is the failure of SLO completion status to predict statistically significant variation in overall self-efficacy and in the sub-constructs of instructional leadership, moral leadership, and managerial leadership, with the exception of the difference in moral leadership between those who met SLO 2 and those who did not meet SLO 2. As it was measured through an independent samples *t*-test and by both parametric and non-parametric techniques, there was found to be a statistically significant effect of SLO 2 passing status on differences in

moral leadership. But the result is the inverse of the hypothesis—as principals failing to attain their SLO 2 showed a higher mean score in this sublevel construct. Further examination will be made in the final section with regard to each question in this category.

Survey Section II: Albert Bandura's Determinants of Self Efficacy

The purpose of this section is to present the results of the second portion of the survey examining the four determinants of self-efficacy as defined by Albert Bandura. The results have been divided into three parts. First, the research questions and hypotheses of the study were restated. Second, the answers to the research questions were presented. Third, the findings of the study were summarized by means a separate evaluation of each of the ten null hypotheses.

To answer Research Questions 3 and 4, the means and standard deviations were calculated for the twelve questions and each subset of the determinants of self-efficacy: enactive mastery, vicarious learning, verbal persuasion, and physiological arousal. Each subset is examined by three survey items.

The questions designed to explore **enactive mastery** are as follows:

While working through the SLO process, to what extent did you...

- modify targets based on formative data?
- document and celebrate incremental successes throughout the process?
- document evidence of support and intervention for your SLOs?

The questions related to **vicarious learning** are the following:

While working through the SLO process, to what extent did you...

- have access to exemplars of SLOs and evaluative rubrics?

- experience personal sharing of your SLOs and progress with your peers?
- experience system sharing of data and lessons learned throughout this process?

The questions related to **verbal persuasion** are the following:

While working through the Student Learning Outcomes process, to what extent did you...

- receive authentic feedback from your evaluator that assisted in setting and monitoring your SLOs?
- receive suggestions of alternate strategies to achieve your SLOs?
- receive encouragement and persistent support from central office staff?

The questions related to **physiological arousal** are the following:

While working through the Student Learning Outcomes process, to what extent did you...

- have direct control over the processes and strategies used to achieve your SLO goals?
- perceive the SLOs as an authentic evaluation of your leadership?
- believe your SLOs aligned to the vision set forth by the school system leadership?

The composite mean score for the twenty-six principals completing the determinants section of the survey was 75.23 out of a possible 108 points. When examining each determinant, out of a possible 27 points, the highest mean score for a determinant was in physiological arousal (M = 20.15), followed by a tie for enactive mastery and verbal persuasion (M = 19.31), with the mean for vicarious learning (M =

16.46) coming in last. These data are depicted in Table 3.

Table 3. *Means, Standard Deviation, and Ranges for Total Determinants of Self-Efficacy and Subscales*

Determinants	Mean	Standard Deviation	Range
Total Scale	75.23	13.85	49–97
Mastery Learning	19.31	3.47	15–23
Vicarious Learning	16.46	5.16	8–26
Verbal Persuasion	19.31	5.45	8–26
Physiological Arousal	20.15	4.12	9–26

Research Question 3 Results: Determinants of Self-Efficacy and SLO 1

The following research questions and hypotheses guided the study for the second set of data collected by the survey sent to participants. The data were analyzed by using the attainment of SLO 1 as the variable and the responses were organized by those who met the SLO 1 compared to those who did not. The twelve questions were grouped according to Bandura's four determinants of self-efficacy: enactive mastery, vicarious learning, verbal persuasion, and physiological arousal. The research questions and hypotheses are as follows:

RQ3.1: Is there a significant difference between mean **effective mastery** scores for (a) principals who met the SLO 1 standard and (b) principals who did not meet the SLO 1 standard?

- H*3.1*$_0$: Mean effective mastery scores for principals who met the SLO 1 standard = Mean effective mastery scores for principals who did not meet the SLO 1 standard.
- H*3.1*$_A$: Mean effective mastery scores for principals who met the SLO 1 standard ≠ Mean effective mastery scores for principals who did not meet the SLO 1 standard.

RQ3.2: Is there a significant difference between mean **vicarious learning** scores for (a) principals who met the SLO 1 standard and (b) principals who did not meet the SLO 1 standard?

- H*3.2*$_0$: Mean vicarious learning scores for principals who met the SLO 1 standard = Mean vicarious learning scores for principals who did not meet the SLO 1 standard.
- H*3.2*$_A$: Mean vicarious learning scores for principals who met the SLO 1 standard ≠ Mean vicarious learning scores for principals who did not meet the SLO 1 standard.

RQ3.3: Is there a significant difference between mean **verbal persuasion** scores for (a) principals who met the SLO 1 standard and (b) principals who did not meet the SLO 1 standard?

- H*3.3*$_0$: Mean verbal persuasion scores for principals who met the SLO 1 standard = Mean verbal persuasion scores for principals who did not meet the SLO 1 standard.

- H*3.3*$_A$: Mean verbal persuasion scores for principals who met the SLO 1 standard ≠ Mean verbal persuasion scores for principals who did not meet the SLO 1 standard.

RQ3.4: Is there a significant difference between mean **physiological arousal** scores for (a) principals who met the SLO 1 standard and (b) principals who did not meet the SLO 1 standard?

- H*3.4*$_0$: Mean physiological arousal scores for principals who met the SLO 1 standard = Mean physiological arousal scores for principals who did not meet the SLO 1 standard.
- H*3.4*$_A$: Mean physiological arousal scores for principals who met the SLO 1 standard ≠ Mean physiological arousal scores for principals who did not meet the SLO 1 standard.

Research Question 3.1 Results: Effective Mastery and SLO 1

Scores on effective mastery were contrasted based on whether principals met or failed to meet the SLO 1 standard. There were eighteen principals who reported meeting the SLO 1 standard and eight principals who reported not meeting the SLO 1 standard. Because the sorting group (meeting or not meeting the SLO 1 standard) was dichotomous and independent, and because the dependent variable of score was continuous, both an independent samples *t*-test and its non-parametric equivalent, the Mann-Whitney U test, were applied.

The independent samples *t*-test indicated that the difference, -0.32, between effective mastery scores for principals passing SLO 1 (M = 19.33, SD = 3.63, SEM = 0.86, n = 18) and effective mastery scores for principals not passing SLO 1 (M = 19.65,

SD = 3.44, *SEM* = 1.22, *n* = 8) was not statistically significant, *t*(24) = -0.21, *p* = 0.84. The 95% confidence interval of the difference between these two groups, -3.46 to 2.82, included the value of 0, providing another means to infer that there was no significant effective mastery difference between principals passing SLO 1 and principals not passing SLO 1. These results were triangulated by the Mann-Whitney U test, *U* = 67.50, *z* = 0.22, *p* = 0.83. Therefore, by both parametric and non-parametric means, it was established that there was no significant difference in the effective mastery scores of principals who passed SLO 1 and principals who did not pass SLO 1.

The box plot in Figure 11 below contains the comparison of mean effective mastery scores for principals who passed SLO 1 and principals who did not pass SLO 1. The box plot clearly indicates the existence of a substantial overlap in the distributions of effective mastery scores for principals who passed SLO 1 and principals who did not pass SLO 1.

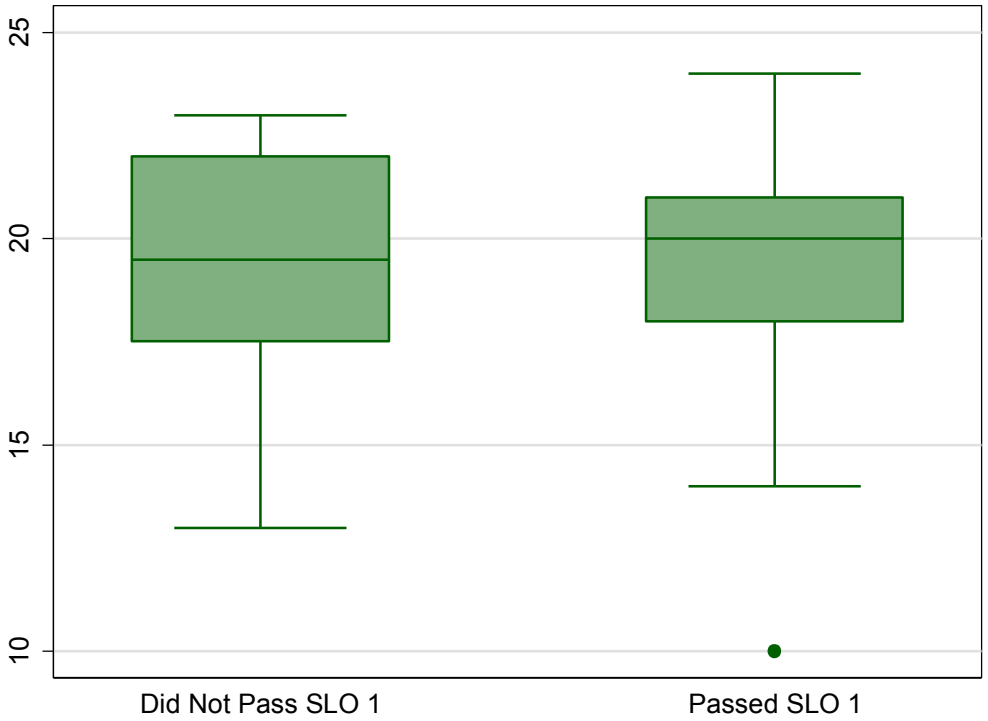

Figure 11. Box plot, comparison of effective mastery by SLO 1 status.
Note: Original graphic generated in Stata / SE 14.2.

The null hypothesis for effective mastery differences between principals who passed SLO 1 and principals who did not pass SLO 1 was that mean effective mastery would be identical between these two groups. On the basis of both the independent samples *t*-test and the Mann-Whitney U test, whose results have been reported above for effective mastery as related to SLO 1 passing status, the null hypothesis could not be rejected. There was insufficient evidence, at an Alpha of 0.05, for the existence of a significant difference in SLO 1 scores between principals who passed SLO 1 and principals who did not pass SLO 1.

Research Question 3.2 Results: Vicarious Learning and SLO 1

Scores on vicarious learning were contrasted based on whether principals met or failed to meet the SLO 1 standard. There were eighteen principals who reported meeting the SLO 1 standard and eight principals who reported not meeting the SLO 1 standard. Because the sorting group (meeting or not meeting the SLO 1 standard) was dichotomous and independent, and because the dependent variable of score was continuous, both an independent samples *t*-test and its non-parametric equivalent, the Mann-Whitney U test, were applied.

The independent samples *t*-test indicated that the difference, 0.51, between vicarious learning scores for principals passing SLO 1 ($M = 16.11$, $SD = 5.60$, $SEM = 1.32$, $n = 18$) and vicarious learning scores for principals not passing SLO 1 ($M = 17.25$, $SD = 4.20$, $SEM = 1.49$, $n = 8$) was not statistically significant, $t(24) = 0.51$, $p = 0.61$. The 95% confidence interval of the difference between these two groups, -5.73 to 3.45, included the value of 0, providing another means to infer that there was no significant vicarious learning difference between principals passing SLO 1 and principals not passing SLO 1. These results were triangulated by the Mann-Whitney U test, $U = 59.50$, $z = -0.67$, $p = 0.50$. Therefore, by both parametric and non-parametric means, it was established that there was no significant difference in the vicarious learning scores of principals who passed SLO 1 and principals who did not pass SLO 1.

The box plot in Figure 12 below contains the comparison of mean vicarious learning scores for principals who passed SLO 1 and principals who did not pass SLO 1. The box plot clearly indicates the existence of a substantial overlap in the distributions of vicarious learning scores for principals who passed SLO 1 and principals who did not

pass SLO 1.

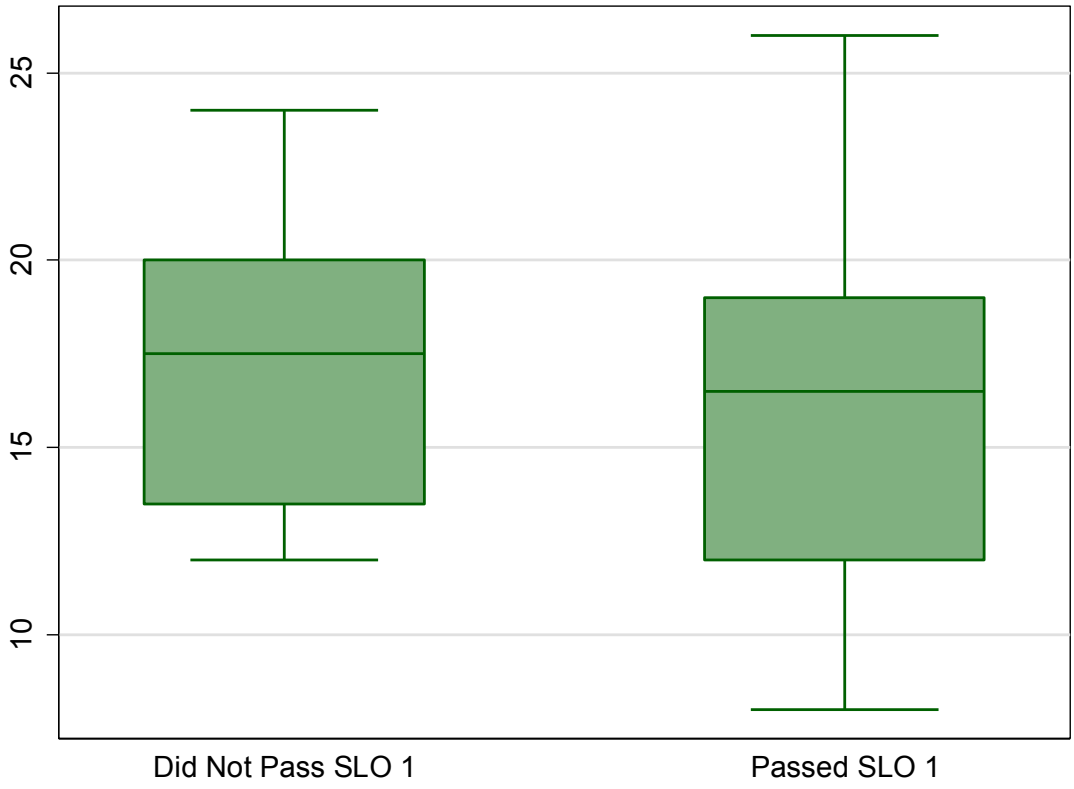

Figure 12. Box plot, comparison of vicarious learning by SLO 1 status.
Note: Original graphic generated in Stata / SE 14.2.

The null hypothesis for vicarious learning differences between principals who passed SLO 1 and principals who did not pass SLO 1 was that mean vicarious learning would be identical between these two groups. On the basis of both the independent samples *t*-test and the Mann-Whitney U test, whose results have been reported above for vicarious learning as related to SLO 1 passing status, the null hypothesis could not be rejected. There was insufficient evidence, at an Alpha of .05, for the existence of a significant difference in SLO 1 scores between principals who passed SLO 1 and principals who did not pass SLO 1.

Research Question 3.3 Results: Verbal Persuasion and SLO 1

Scores on verbal persuasion were contrasted based on whether principals met or failed to meet the SLO 1 standard. There were eighteen principals who reported meeting the SLO 1 standard and eight principals who reported not meeting the SLO 1 standard. Because the sorting group (meeting or not meeting the SLO 1 standard) was dichotomous and independent, and because the dependent variable of score was continuous, both an independent samples *t*-test and its non-parametric equivalent, the Mann-Whitney U test, were applied.

The independent samples *t*-test indicated that the difference, 0.81, between verbal persuasion scores for principals passing SLO 1 ($M = 19.56$, $SD = 4.67$, $SEM = 1.10$, $n = 18$) and verbal persuasion scores for principals not passing SLO 1 ($M = 18.75$, $SD = 7.25$, $SEM = 2.56$, $n = 8$) was not statistically significant, $t(24) = 0.34$, $p = 0.73$. The 95% confidence interval of the difference between these two groups, -4.05 to 5.68, included the value of 0, providing another means to infer that there was no significant verbal persuasion difference between principals passing SLO 1 and principals not passing SLO 1. These results were triangulated by the Mann-Whitney U test, $U = 71.00$, $z = -0.03$, $p = 0.98$. Therefore, by both parametric and non-parametric means, it was established that there was no significant difference in the verbal persuasion scores of principals who passed SLO 1 and principals who did not pass SLO 1.

The box plot in Figure 13 below contains the comparison of mean verbal persuasion scores for principals who passed SLO 1 and principals who did not pass SLO 1. The box plot clearly indicates the existence of a substantial overlap in the distributions

of verbal persuasion scores for principals who passed SLO 1 and principals who did not pass SLO 1.

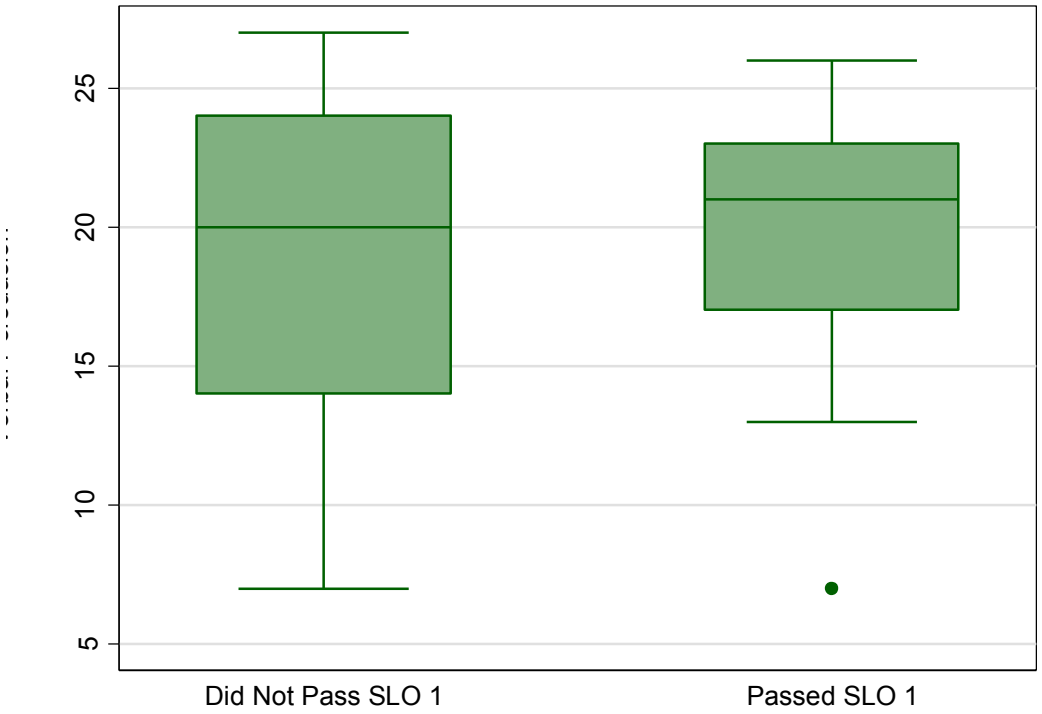

Figure 13. Box plot, comparison of verbal persuasion by SLO 1 status.
Note: Original graphic generated in Stata / SE 14.2.

The null hypothesis for verbal persuasion differences between principals who passed SLO 1 and principals who did not pass SLO 1 was that mean verbal persuasion would be identical between these two groups. On the basis of both the independent samples *t*-test and the Mann-Whitney U test, whose results have been reported above for verbal persuasion as related to SLO 1 passing status, the null hypothesis could not be rejected. There was insufficient evidence, at an Alpha of 0.05, for the existence of a significant difference in SLO 1 scores between principals who passed SLO 1 and principals who did not pass SLO 1.

Research Question 3.4 Results: Physiological Arousal and SLO 1

Scores on physiological arousal were contrasted based on whether principals met or failed to meet the SLO 1 standard. There were eighteen principals who reported meeting the SLO 1 standard and eight principals who reported not meeting the SLO 1 standard. Because the sorting group (meeting or not meeting the SLO 1 standard) was dichotomous and independent, and because the dependent variable of score was continuous, both an independent samples *t*-test and its non-parametric equivalent, the Mann-Whitney U test, were applied.

The independent samples *t*-test indicated that the difference, -0.40, between physiological arousal scores for principals passing SLO 1 (M = 20.28, SD = 3.83, SEM = 0.90, n = 18) and physiological arousal scores for principals not passing SLO 1 (M = 19.88, SD = 4.97, SEM = 1.76, n = 8) was not statistically significant, $t(24) = -0.23$, p = 0.82. The 95% confidence interval of the difference between these two groups, -4.08 to 3.28, included the value of 0, providing another means to infer that there was no significant physiological arousal difference between principals passing SLO 1 and principals not passing SLO 1. These results were triangulated by the Mann-Whitney U test, U = 69.50, z = -0.11, p = 0.91. Therefore, by both parametric and non-parametric means, it was established that there was no significant difference in the physiological arousal scores of principals who passed SLO 1 and principals who did not pass SLO 1.

The box plot in Figure 14 below contains the comparison of mean physiological arousal scores for principals who passed SLO 1 and principals who did not pass SLO 1. The box plot clearly indicates the existence of a substantial overlap in the distributions of

physiological arousal scores for principals who passed SLO 1 and principals who did not pass SLO 1.

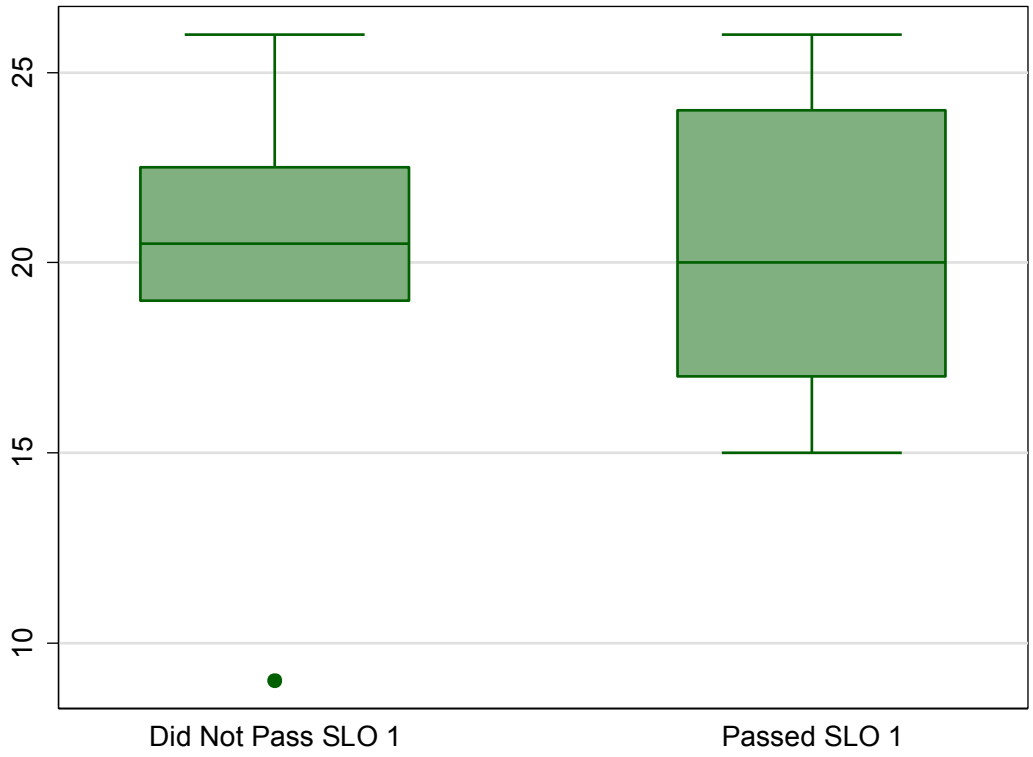

Figure 14. Box plot, comparison of physiological arousal by SLO 1 status.
Note: Original graphic generated in Stata / SE 14.2.

The null hypothesis for physiological arousal differences between principals who passed SLO 1 and principals who did not pass SLO 1 was that mean physiological arousal would be identical between these two groups. On the basis of both the independent samples *t*-test and the Mann-Whitney U test, whose results have been reported above for physiological arousal as related to SLO 1 passing status, the null hypothesis could not be rejected. There was insufficient evidence, at an Alpha of 0.05, for the existence of a significant difference in SLO 1 scores between principals who passed SLO 1 and principals who did not pass SLO 1.

Research Question 4 Results: Determinants of Self Efficacy and SLO 2

The following research questions and hypotheses guided the study for the final research question using the set of data collected by the survey sent to participants. The data were analyzed by using the attainment of SLO 2 as the variable and the responses were organized by those who met the SLO compared to those who did not. The twelve questions were grouped according to Bandura's four determinants of self-efficacy: enactive mastery, vicarious learning, verbal persuasion, and physiological arousal. The research questions and hypotheses are as follows:

RQ4.1: Is there a significant difference between mean **effective mastery** scores for (a) principals who met the SLO 2 standard and (b) principals who did not meet the SLO 2 standard?

- H*4.1*$_0$: Mean effective mastery scores for principals who met the SLO 2 standard = Mean effective mastery scores for principals who did not meet the SLO 2 standard.
- H*4.1*$_A$: Mean effective mastery scores for principals who met the SLO 2 standard ≠ Mean effective mastery scores for principals who did not meet the SLO 2 standard.

RQ4.2: Is there a significant difference between mean **vicarious learning** scores for (a) principals who met the SLO 2 standard and (b) principals who did not meet the SLO 2 standard?

- H*4.2*$_0$: Mean vicarious learning scores for principals who met the SLO 2 standard = Mean vicarious learning scores for principals who did not meet the SLO 2 standard.
- H*4.2*$_A$: Mean vicarious learning scores for principals who met the SLO 2 standard ≠ Mean vicarious learning scores for principals who did not meet the SLO 2 standard.

RQ4.3: Is there a significant difference between mean **verbal persuasion** scores for (a) principals who met the SLO 2 standard and (b) principals who did not meet the SLO 2 standard?

- H*4.3*$_0$: Mean verbal persuasion scores for principals who met the SLO 2 standard = Mean verbal persuasion scores for principals who did not meet the SLO 2 standard.
- H*4.3*$_A$: Mean verbal persuasion scores for principals who met the SLO 2 standard ≠ Mean verbal persuasion scores for principals who did not meet the SLO 2 standard.

RQ4.4: Is there a significant difference between mean **physiological arousal** scores for (a) principals who met the SLO 2 standard and (b) principals who did not meet the SLO 2 standard?

- H*4.4*$_0$: Mean physiological arousal scores for principals who met the SLO 2 standard = Mean physiological arousal scores for principals who did not meet the SLO 2 standard.

- H*4.4*$_A$: Mean physiological arousal scores for principals who met the SLO 2 standard ≠ Mean physiological arousal scores for principals who did not meet the SLO 2 standard.

For the purposes of hypothesis testing, the level of statistical significance of the study was 0.05. All data analysis and graph construction for the study took place in Stata / SE 14.2 software.

Research Question 4.1 Results: Enactive Mastery and SLO 1

Scores on effective mastery were contrasted based on whether principals met or failed to meet the SLO 2 standard. There were twenty principals who reported meeting the SLO 2 standard and six principals who reported not meeting the SLO 2 standard. Because the sorting group (meeting or not meeting the SLO 2 standard) was dichotomous and independent, and because the dependent variable of score was continuous, both an independent samples *t*-test and its non-parametric equivalent, the Mann-Whitney U test, were applied.

The independent samples *t*-test indicated that the difference, -0.76, between effective mastery scores for principals passing SLO 2 (M = 19.45, SD = 3.65, SEM = 0.82, n = 20) and effective mastery scores for principals not passing SLO 2 (M = 20.21, SD = 2.50, SEM = 1.02, n = 6) was not statistically significant, $t(24) = 0.47$, $p = 0.64$. The 95% confidence interval of the difference between these two groups, -4.07 to 2.55, included the value of 0, providing another means to infer that there was no significant effective mastery difference between principals passing SLO 2 and principals not passing SLO 2. These results were triangulated by the Mann-Whitney U test, U = 46.00, z = 0.82,

$p = 0.41$. Therefore, by both parametric and non-parametric means, it was established that there was no significant difference in the effective mastery scores of principals who passed SLO 2 and principals who did not pass SLO 2.

The box plot in Figure 15 below contains the comparison of mean effective mastery scores for principals who passed SLO 2 and principals who did not pass SLO 2. The box plot clearly indicates the existence of a substantial overlap in the distributions of effective mastery scores for principals who passed SLO 2 and principals who did not pass SLO 2.

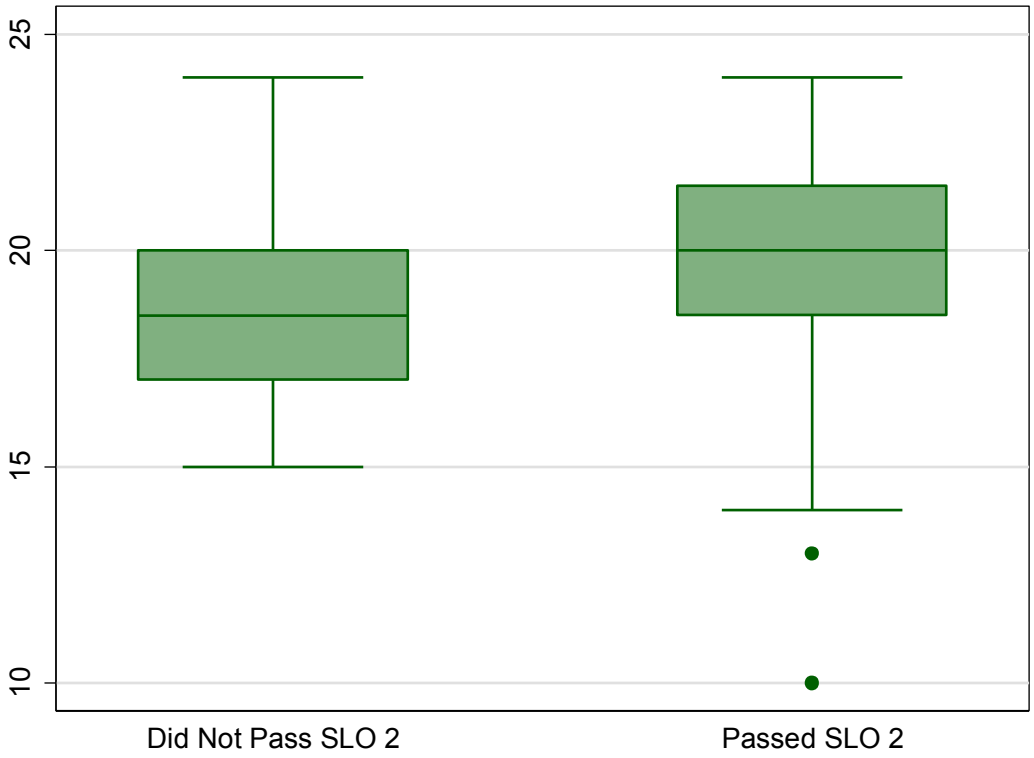

Figure 15. Box plot, comparison of effective mastery by SLO 2 status.
Note: Original graphic generated in Stata / SE 14.2.

The null hypothesis for effective mastery differences between principals who passed SLO 2 and principals who did not pass SLO 2 was that mean effective mastery would be identical between these two groups. On the basis of both the independent

samples *t*-test and the Mann-Whitney U test, whose results have been reported above for effective mastery as related to SLO 2 passing status, the null hypothesis could not be rejected. There was insufficient evidence, at an Alpha of 0.05, for the existence of a significant difference in SLO 2 scores between principals who passed SLO 2 and principals who did not pass SLO 2.

Research Question 4.2 Results: Vicarious Learning and SLO 2

Scores on vicarious learning were contrasted based on whether principals met or failed to meet the SLO 2 standard. There were twenty principals who reported meeting the SLO 2 standard and six principals who reported not meeting the SLO 2 standard. Because the sorting group (meeting or not meeting the SLO 2 standard) was dichotomous and independent, and because the dependent variable of score was continuous, both an independent samples *t*-test and its non-parametric equivalent, the Mann-Whitney U test, were applied.

The independent samples *t*-test indicated that the difference, -1.05, between vicarious learning scores for principals passing SLO 2 ($M = 16.35$, $SD = 5.11$, $SEM = 1.14$, $n = 20$) and vicarious learning scores for principals not passing SLO 2 ($M = 17.40$, $SD = 5.13$, $SEM = 2.09$, $n = 6$) was not statistically significant, $t(24) = 0.44$, $p = 0.66$. The 95% confidence interval of the difference between these two groups, -5.96 to 3.86, included the value of 0, providing another means to infer that there was no significant vicarious learning difference between principals passing SLO 2 and principals not passing SLO 2. These results were triangulated by the Mann-Whitney U test, $U = 56.50$, $z = -0.18$, $p = 0.86$. Therefore, by both parametric and non-parametric means, it was established that there was no significant difference in the vicarious learning scores of

principals who passed SLO 2 and principals who did not pass SLO 2.

The box plot in Figure 16 below contains the comparison of mean vicarious learning scores for principals who passed SLO 2 and principals who did not pass SLO 2. The box plot clearly indicates the existence of a substantial overlap in the distributions of vicarious learning scores for principals who passed SLO 2 and principals who did not pass SLO 2.

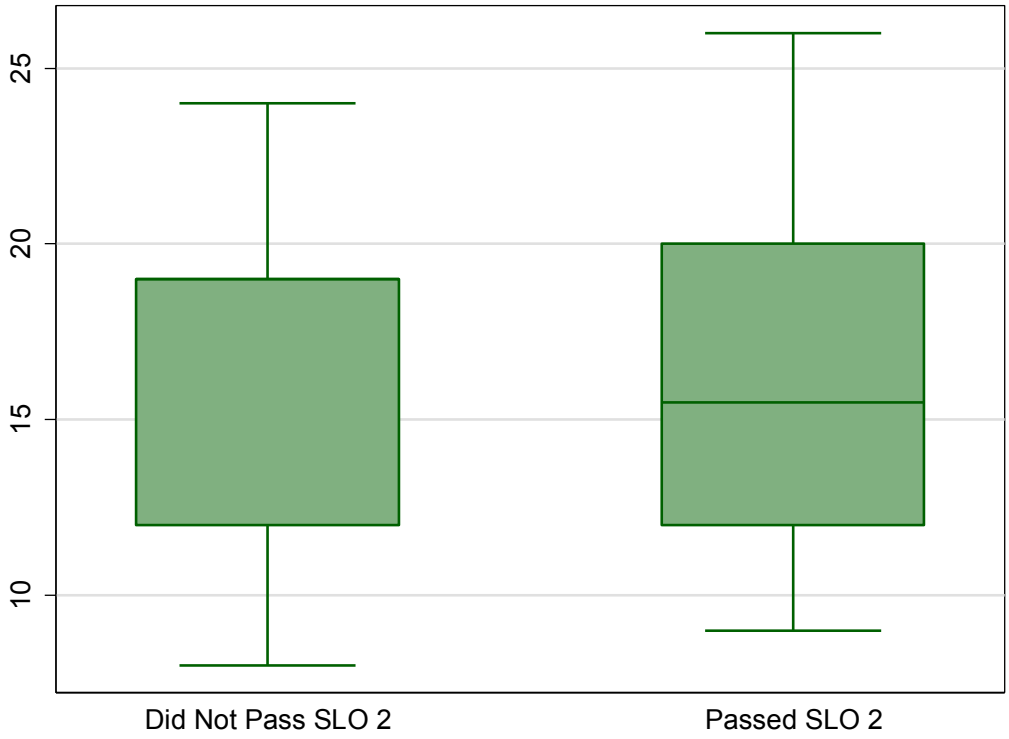

Figure 16. Box plot, comparison of vicarious learning by SLO 2 status.
Note: Original graphic generated in Stata / SE 14.2.

The null hypothesis for vicarious learning differences between principals who passed SLO 2 and principals who did not pass SLO 2 was that mean vicarious learning would be identical between these two groups. On the basis of both the independent samples *t*-test and the Mann-Whitney U test, whose results have been reported above for

vicarious learning as related to SLO 2 passing status, the null hypothesis could not be rejected. There was insufficient evidence, at an Alpha of 0.05, for the existence of a significant difference in SLO 2 scores between principals who passed SLO 2 and principals who did not pass SLO 2.

Research Question 4.3 Results: Verbal Persuasion and SLO 2

Scores on verbal persuasion were contrasted based on whether principals met or failed to meet the SLO 2 standard. There were twenty principals who reported meeting the SLO 2 standard and six principals who reported not meeting the SLO 2 standard. Because the sorting group (meeting or not meeting the SLO 2 standard) was dichotomous and independent, and because the dependent variable of score was continuous, both an independent samples *t*-test and its non-parametric equivalent, the Mann-Whitney U test, were applied.

The independent samples *t*-test indicated that the difference, 0.17, between verbal persuasion scores for principals passing SLO 2 (M = 19.31, SD = 5.45, SEM = 1.22, n = 20) and verbal persuasion scores for principals not passing SLO 2 (M = 18.87, SD = 5.70, SEM = 2.33, n = 6) was not statistically significant, $t(24) = 0.17$, $p = 0.87$. The 95% confidence interval of the difference between these two groups, -4.84 to 5.73, included the value of 0, providing another means to infer that there was no significant verbal persuasion difference between principals passing SLO 2 and principals not passing SLO 2. These results were triangulated by the Mann-Whitney U test, $U = 57.50$, $z = 0.12$, p = 0.90. Therefore, by both parametric and non-parametric means, it was established that there was no significant difference in the verbal persuasion scores of principals who

passed SLO 2 and principals who did not pass SLO 2.

The box plot in Figure 17 below contains the comparison of mean verbal persuasion scores for principals who passed SLO 2 and principals who did not pass SLO 2. The box plot clearly indicates the existence of a substantial overlap in the distributions of verbal persuasion scores for principals who passed SLO 2 and principals who did not pass SLO 2.

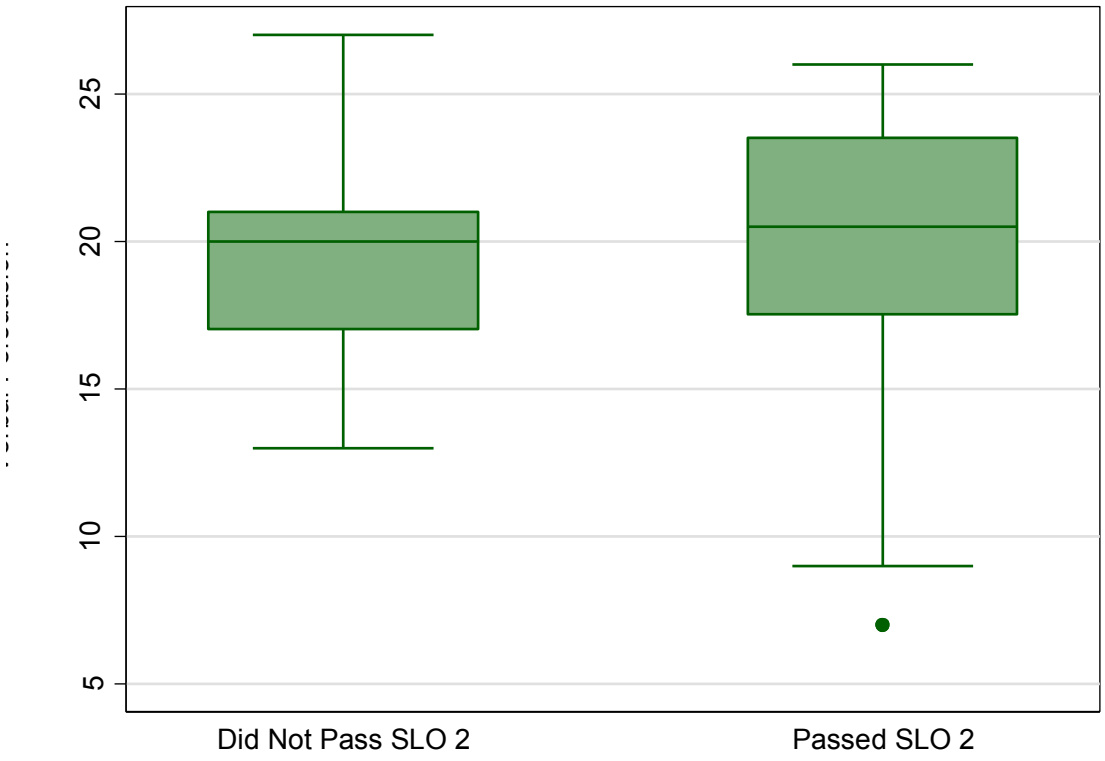

Figure 17. Box plot, comparison of verbal persuasion by SLO 2 status.
Note: Original graphic generated in Stata / SE 14.2.

The null hypothesis for verbal persuasion differences between principals who passed SLO 2 and principals who did not pass SLO 2 was that mean verbal persuasion would be identical between these two groups. On the basis of both the independent samples *t*-test and the Mann-Whitney U test, whose results have been reported above for

verbal persuasion as related to SLO 2 passing status, the null hypothesis could not be rejected. There was insufficient evidence, at an Alpha of 0.05, for the existence of a significant difference in SLO 2 scores between principals who passed SLO 2 and principals who did not pass SLO 2.

Research Question 4.4 Results: Physiological Arousal and SLO 2

Scores on physiological arousal were contrasted based on whether principals met or failed to meet the SLO 2 standard. There were twenty principals who reported meeting the SLO 2 standard and six principals who reported not meeting the SLO 2 standard. Because the sorting group (meeting or not meeting the SLO 2 standard) was dichotomous and independent, and because the dependent variable of score was continuous, both an independent samples *t*-test and its non-parametric equivalent, the Mann-Whitney U test, were applied.

The independent samples *t*-test indicated that the difference, 3.05, between physiological arousal scores for principals passing SLO 2 (M = 19.45, SD = 4.12, SEM = 0.92, n = 20) and physiological arousal scores for principals not passing SLO 2 (M = 22.50, SD = 3.39, SEM = 1.38, n = 6) was not statistically significant, $t(24) = 1.64$, p = 0.11. The 95% confidence interval of the difference between these two groups, -0.77 to 6.87, included the value of 0, providing another means to infer that there was no significant physiological arousal difference between principals passing SLO 2 and principals not passing SLO 2. These results were triangulated by the Mann-Whitney U test, U = 33.50, z = -1.58, p = .11. Therefore, by both parametric and non-parametric means, it was established that there was no significant difference in the physiological

arousal scores of principals who passed SLO 2 and principals who did not pass SLO

The box plot in Figure 18 below contains the comparison of mean physiological arousal scores for principals who passed SLO 2 and principals who did not pass SLO 2. The box plot clearly indicates the existence of a substantial overlap in the distributions of physiological arousal scores for principals who passed SLO 2 and principals who did not pass SLO 2.

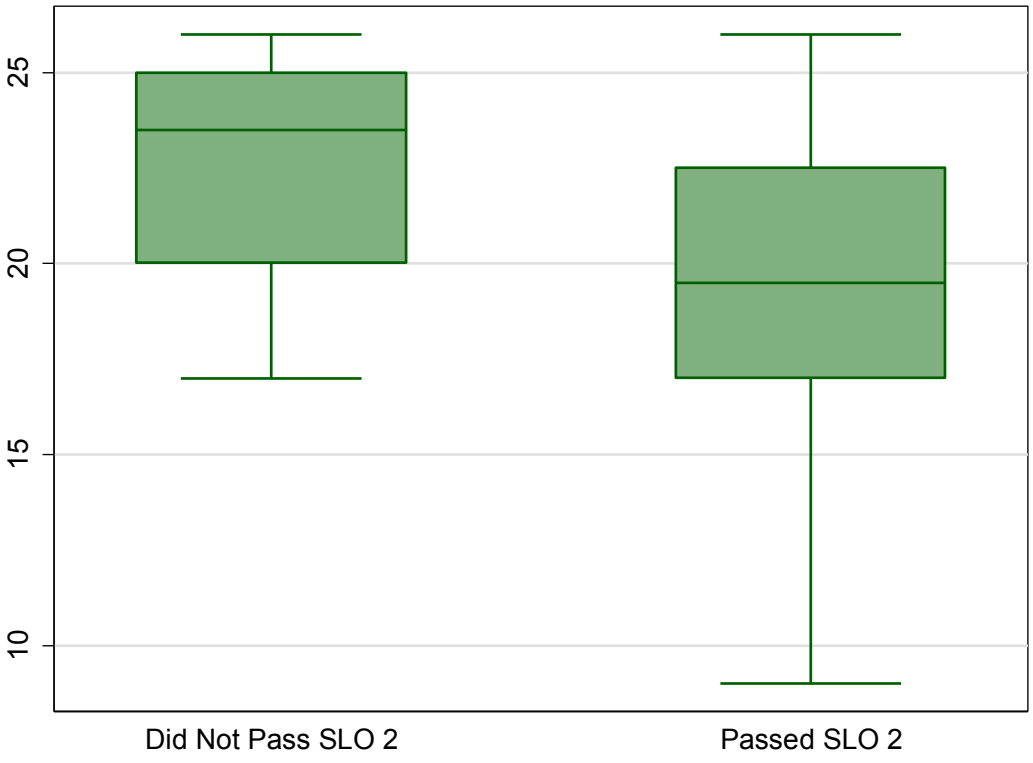

Figure 18. Box plot, comparison of physiological arousal by SLO 2 status.
Note: Original graphic generated in Stata / SE 14.2.

The null hypothesis for physiological arousal differences between principals who passed SLO 2 and principals who did not pass SLO 2 was that mean physiological arousal would be identical between these two groups. On the basis of both the

independent samples *t*-test and the Mann-Whitney U test, whose results have been reported above for physiological arousal as related to SLO 2 passing status, the null hypothesis could not be rejected. There was insufficient evidence, at an Alpha of 0.05, for the existence of a significant difference in SLO 2 scores between principals who passed SLO 2 and principals who did not pass SLO 2.

Discussion and Further Examination of Survey Results by Radar Graph

Principals' Self-Efficacy Survey

While there was no statistical significance in the findings from the Principals' Self-Efficacy Survey, it is worthwhile to examine the variance in mean scores for the individual questions between the principals who attained the SLOs and those who did not. This can be visually represented using a radar graph with each question as a point and overlaying the principals who met their SLO 1 for student achievement (*n* of 18) compared to those who did not (*n* of 8). When this is done for the first sublevel construct from the PSES, *instructional leadership*, it yields the following Figure 19.

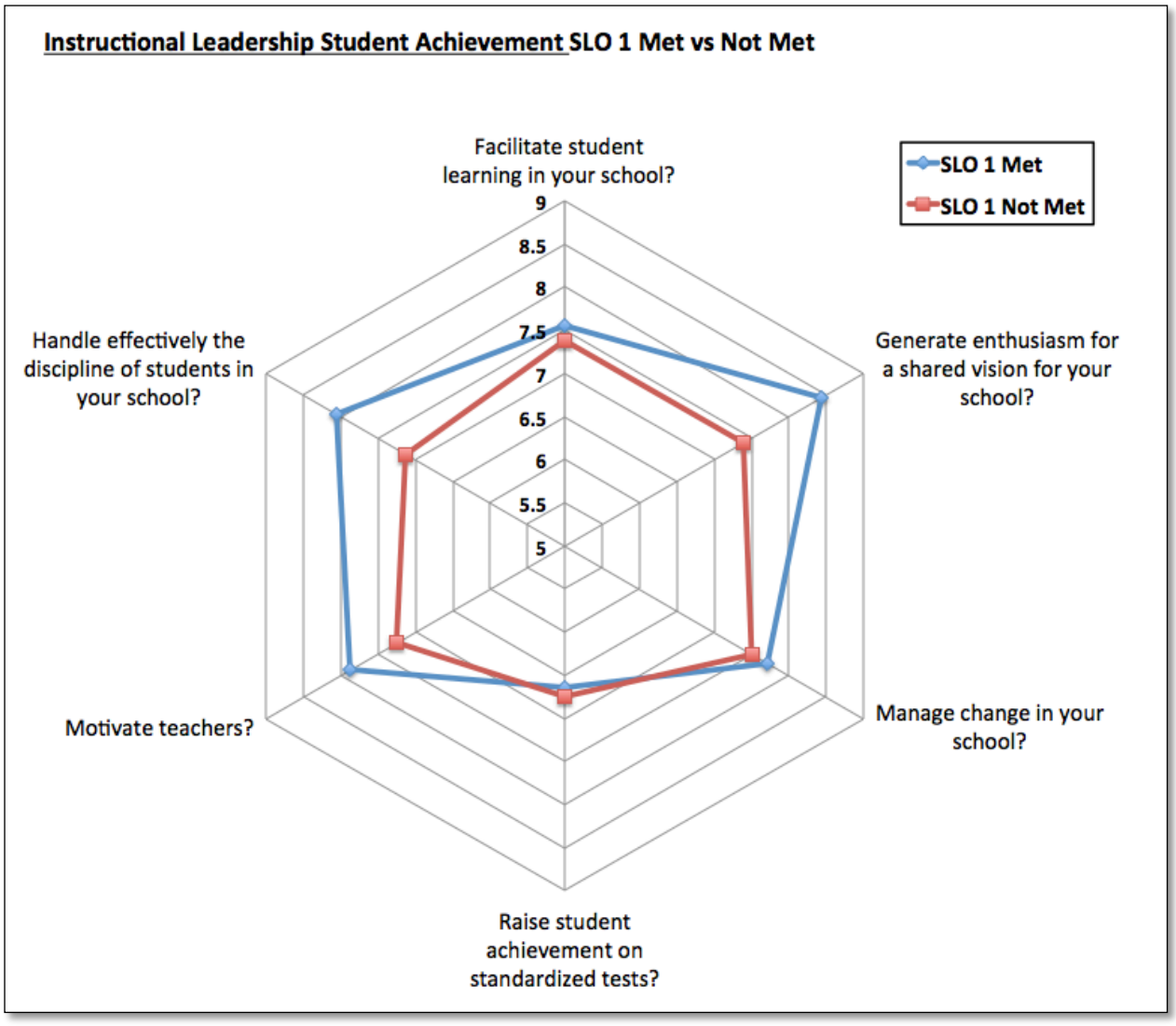

Figure 19. Radar graph, comparison of instructional leadership by SLO 1 status.
Note: Original graphic generated in Microsoft Excel.

The results for this sublevel construct, instructional leadership, follows the hypothesis that principals attaining the SLO for student achievement would have a higher mean score for each question. The only outlier is the question regarding "raising student achievement on standardized tests"; the principals who did not meet the SLO 1 show a slightly higher mean score of 6.75 versus 6.65 for those who met their SLO 1. Within the confines of the study, there is no further opportunity for analysis, but the data do draw

attention to the apparent disconnect between the principals' perception of their ability and the actual attainment of the student achievement outcome.

Figure 20 uses the same visual and overlays the data for those principals who met the SLO 2 for culture and climate (*n* of 20) versus those who did not (*n* of 8).

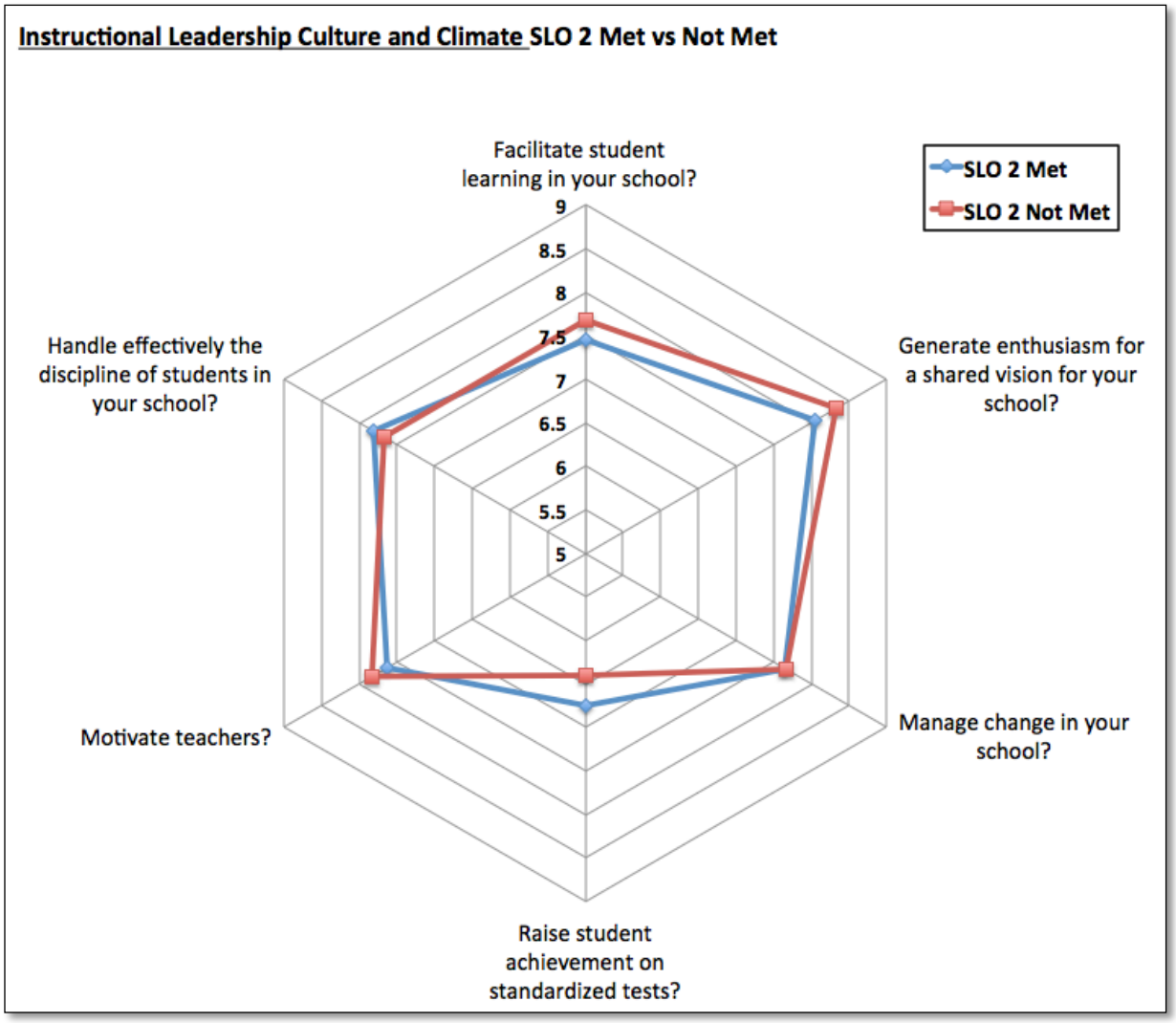

Figure 20. Radar graph, comparison of instructional leadership by SLO 1 status. *Note*: Original graphic generated in Microsoft Excel.

The data here do not follow the hypothesis that principals meeting their SLO 2 for climate and culture will have a higher mean score than those who did not. For this sublevel construct for instructional leadership, principals not meeting their SLO 2 posted

a higher mean score for “facilitate student learning in your school” with a mean of 7.76 for those not meeting SLO 2 versus a mean of 7.45 for those who did. Principals not meeting SLO 2 also posted a higher mean score for “motivate teachers” with a mean of 7.83 for those not meeting SLO 2 versus 7.63 for those who did. The final outlier is for “generate enthusiasm for a shared vision for your school” with a mean score of 8.33 for those who did not meet their SLO 2 versus 8.05 for those who did. Again, limitations with the study do not allow further investigation, but it is evident that the principals not meeting their student achievement SLO have a higher perception of their ability to motivate, facilitate, and generate enthusiasm in their schools then those actually meeting.

Moving to the next PSES sublevel construct of moral leadership, the data for those meeting the SLO 1 for student achievement (*n* of 18) compared to those who did not (*n* of 8) are represented by Figure 21.

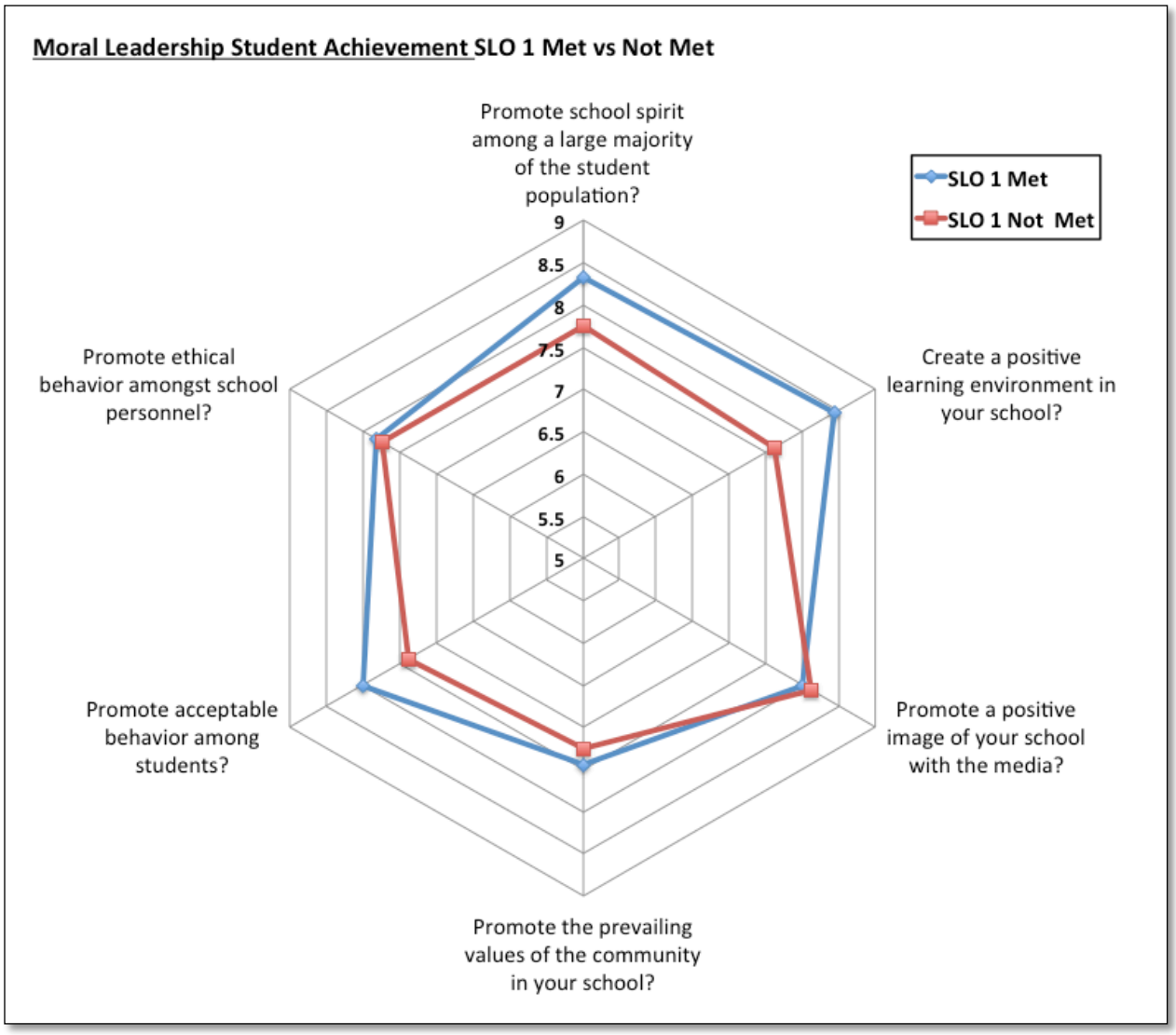

Figure 21. Radar graph, comparison of moral leadership by SLO 1 status.
Note: Original graphic generated in Microsoft Excel.

In this representation, the results align to the hypothesis that principals meeting their SLO 1 for student achievement will score a higher mean score for each question with the exception of the question of "promote a positive image of your school in the media" with a mean score of 8.13 for those who did not meet their SLO 1 versus 8.0 for those who did.

For climate and culture SLO 2, the data are represented in Figure 22. The data for this sublevel construct do not support the hypothesis that the mean score for those

meeting this SLO 2 will be higher than those who did not. The only question that had those principals who met SLO 2 (*n* of 20) scoring higher that those who did not (*n* of 6) is for "create a positive learning environment in your school" with a mean score of 8.2 for those who met their SLO 2 versus 8.17 for those who did not.

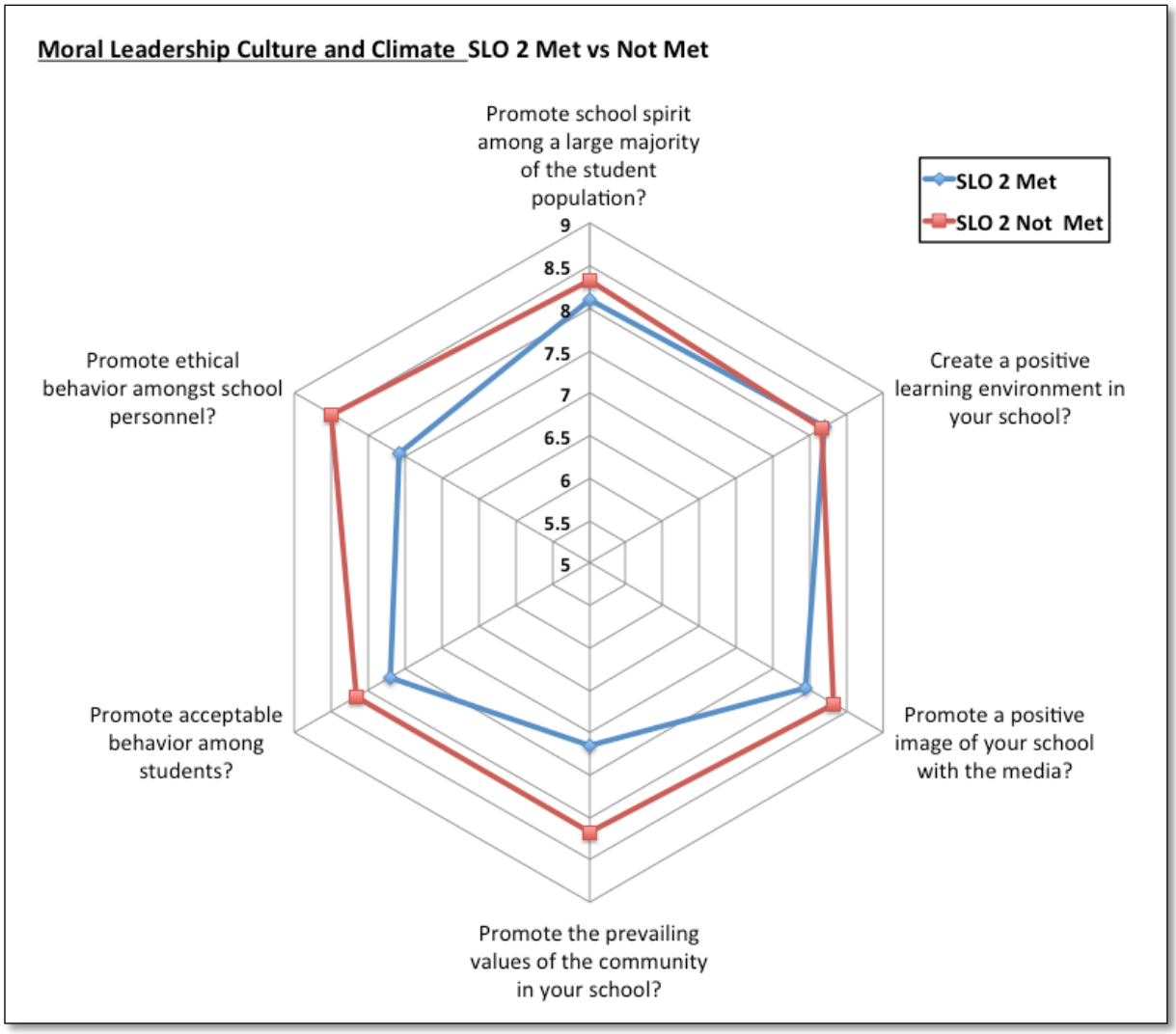

Figure 22. Radar graph, comparison of moral leadership by SLO 2 status.
Note: Original graphic generated in Microsoft Excel.

This was the only sublevel construct that yielded a statistically significant result when run through a t-test. The mean moral leadership for those who met SLO 2 ($M = 46.30$, $SD = 4.03$, $n = 20$) was compared to the mean moral leadership for those who did not meet

SLO 2 (M = 49.67 SD = 2.50, n =6). The calculated difference of 3.37 (95% confidence interval = -0.25 to 6.98) was found to be statistically significant, such that the mean moral leadership of those who did not pass SLO 2 was higher than the mean moral leadership of those who passed SLO 2, $t(24) = 1.92$, $p = 0.03$. The Mann-Whitney U test for the difference in moral leadership by mean SLO 2 pass status was also significant, $MWU = 29$, $z = -1.86$, $p = 0.03$. Therefore, by both parametric and non-parametric techniques, there was found to be a statistically significant effect of SLO 2 passing status on differences in moral leadership. While this is statistically significant, it is the inverse of the hypothesis and warrants further study. The sample size of respondents and possible bias of their answers should be considered before reaching any conclusions.

The final sublevel construct of the PSES is managerial leadership. When the mean scores are placed in the radar graph, Figure 23, the hypothesis that those principals meeting their SLO 1 for student achievement (n of 18) will have a higher mean than those who did not (n of 8) is supported by data collected except for "maintain control of your own daily schedule" with a slightly higher mean for those not meeting this SLO 1, 6.5 versus 6.39 for those who did and the dramatic variance in the mean scores for "shape the operational policies and procedures that are necessary to manage school." A mean score of 6.75 was recorded for the eight principals who did not meet their SLO 1 compared to a mean score of 5.72 for those who did.

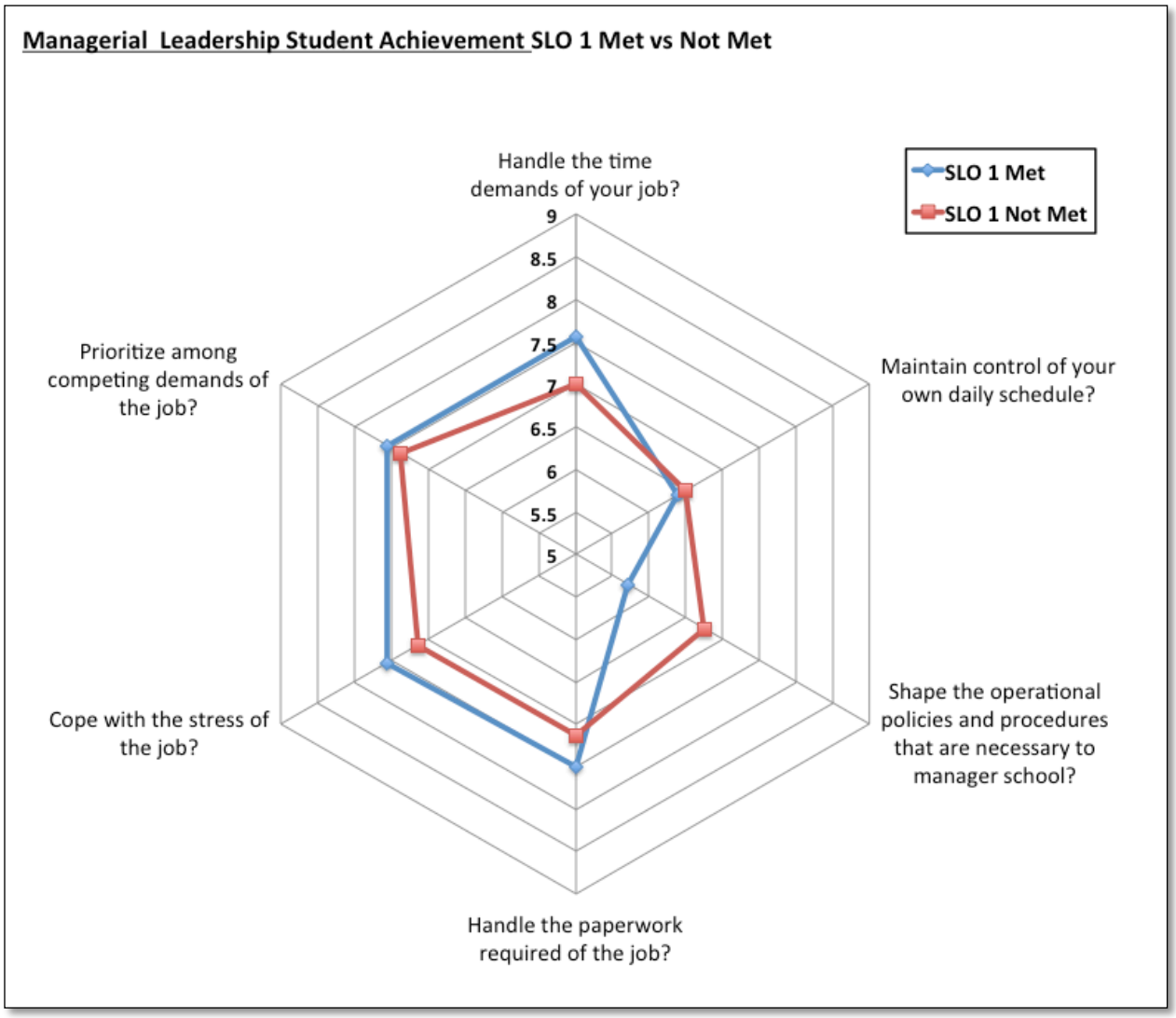

Figure 23. Radar graph, comparison of managerial leadership by SLO 1 status.
Note: Original graphic generated in Microsoft Excel.

The last radar graph in the PSES section, Figure 24, compares the mean scores for the principals who met their SLO 2 for culture and climate (*n* of 20) versus those who did not (*n* of 6). For this sublevel construct, the hypothesis is only supported by the data "handle the paperwork for the job" with a mean score of 7.5 for those who met the SLO 2 versus 7.0 for those who did not, with a mean score of 7.4 for "handle the time demands of your job" versus a slightly lower 7.33 for those who did not.

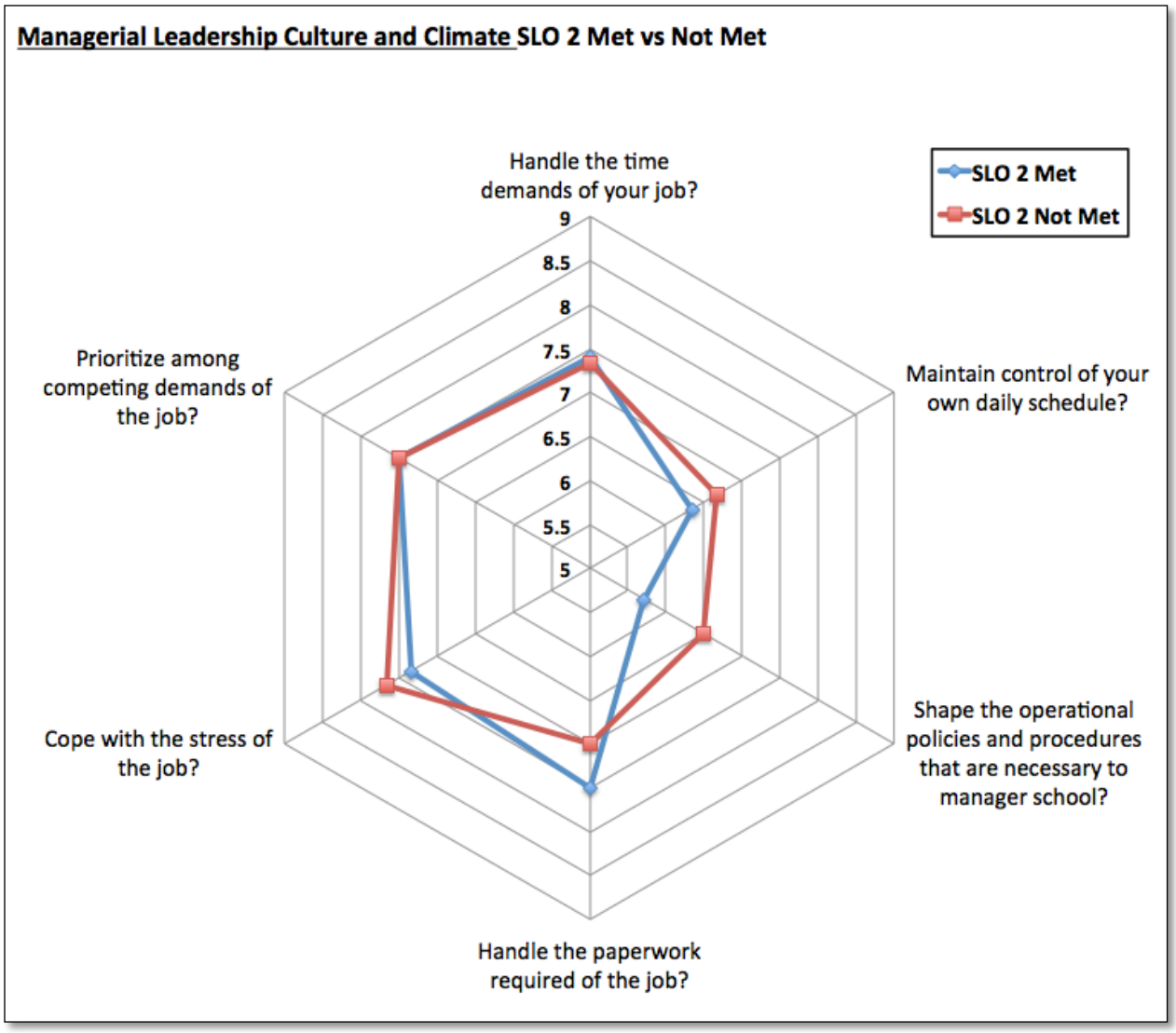

Figure 24. Radar graph, comparison of managerial leadership by SLO 2 status.
Note: Original graphic generated in Microsoft Excel.

Bandura's Determinants of Self-Efficacy

While there was not statistical significance in the results comparing principals who met their SLOs to those who did not in the mean scores on the determinants of self-efficacy as defined by Albert Badura, it is worthwhile to examine the variance on individual questions. Again, this is visually represented using a radar graph with each question as a point and overlaying the principals who

met their SLO 1 for student achievement (*n* of 18) compared to those who did not (*n* of 8). See Figure 25.

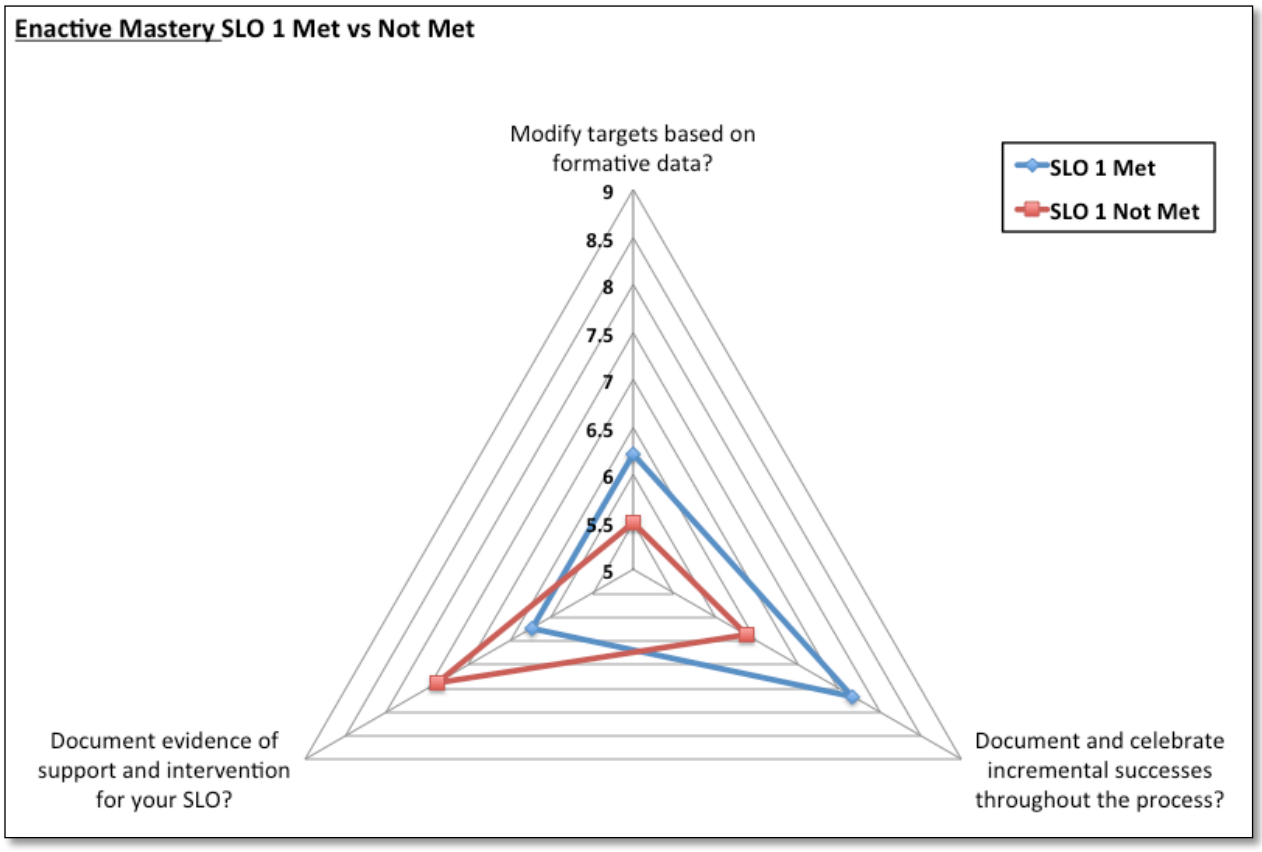

Figure 25. Radar graph, comparison of enactive mastery by SLO 1 status.
Note: Original graphic generated in Microsoft Excel.

Figure 26 shows that the responses for the three questions regarding enactive mastery had a higher mean score for those who met their student achievement SLO, with the exception of the question of whether or not the principals "documented evidence of support and intervention for their SLO?" This question had a higher mean score of 7.38 versus a mean of 6.22 for those who did not met their student SLO. This would make sense, as principals who did not meet their target would have sought to understand why and would focus on compiling evidence as to what had

been done to demonstrate and document their attempted support and intervene when they say that performance was not meeting goals.

Applying this same visual to the data from the questions for enactive mastery relative to SLO 2 culture and climate and the principals who met their SLO (*n* of 20) versus those who did not (*n* of 6) yields the following Figure 25.

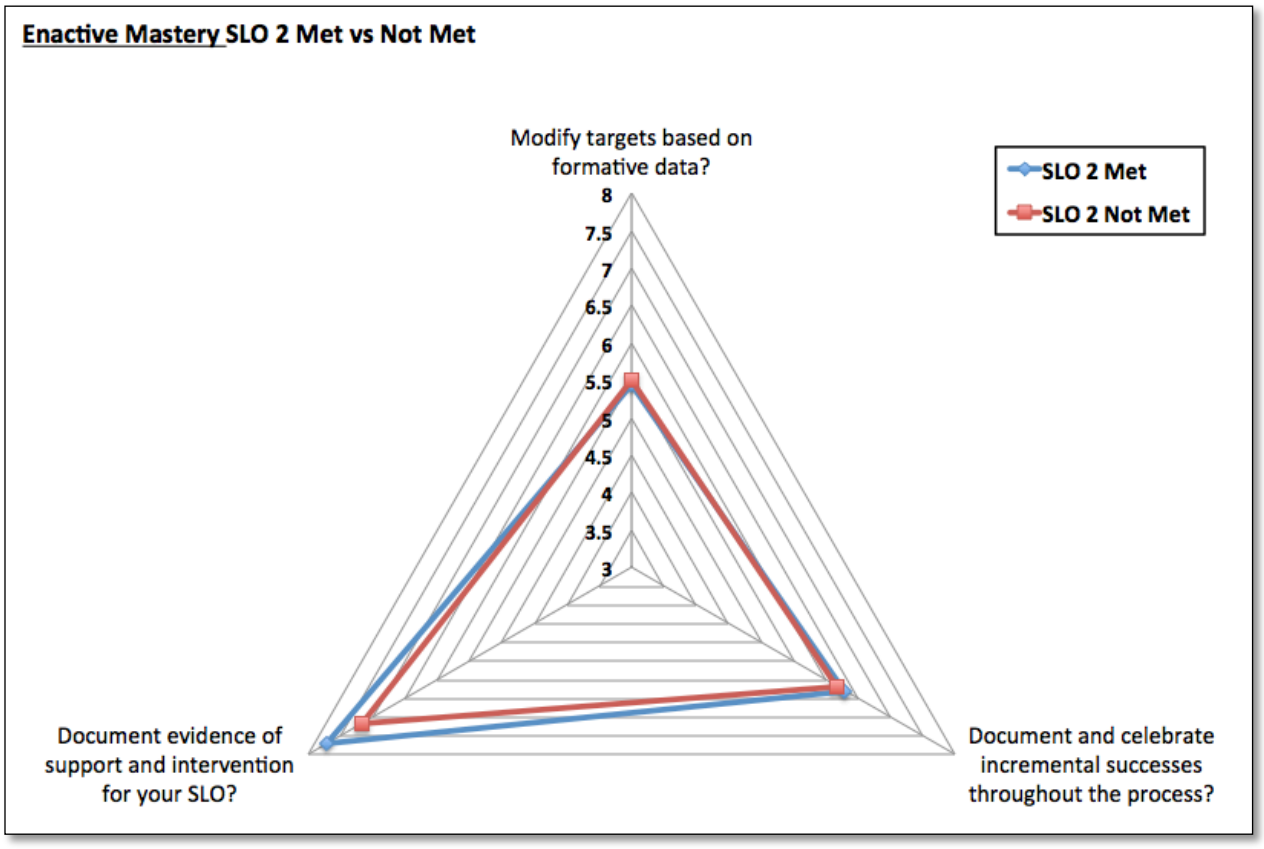

Figure 26. Radar graph, comparison of enactive mastery by SLO 2 status.
Note: Original graphic generated in Microsoft Excel.

In this, one can see that the mean for each question was quite similar with the highest scores being attributed to the question regarding the evidence of support and intervention provided.

The radar graph for vicarious learning for SLO 1 shows the mean score for all principals

varied little with the exception of the "experience of sharing of your SLOs and progress with your peers."

Moving to the second determinant of self-efficacy, vicarious learning, the radar graph, Figure 27, compares the scores for those meeting SLO 1 for student achievement (*n* of 18) compared to those who did not (*n* of 8).

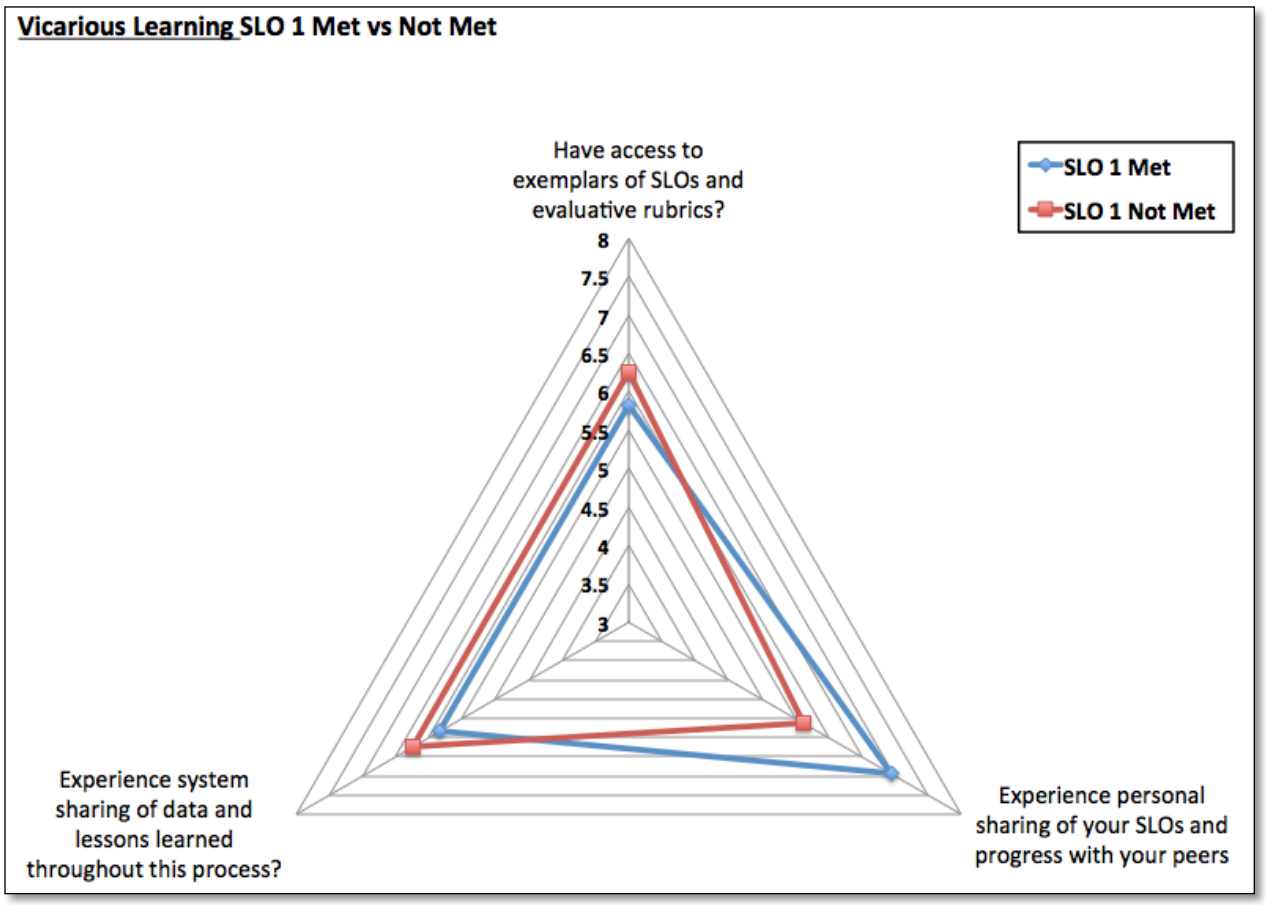

Figure 27. Radar graph, comparison of vicarious learning by SLO 1 status.
Note: Original graphic generated in Microsoft Excel.

The data do not support the hypothesis that principals meeting with SLO 1 would have a higher mean score for two of the three questions. "Have access to exemplars of SLOs and evaluative rubrics" had a mean score of 6.25 for those who did not meet with SLO 1 versus 5.83 for those who did. Principals who did not meet their SLO had a mean score of 6.25 for "experience system sharing of data and lessons learned throughout the

process" versus 5.83 for those who did. It could be inferred from these data that principals who struggled to meet their goals were provided exemplars and had opportunities to share their experiences from the process even as this did not result in their attaining their goals. The final data point relative to "experience personal sharing of your SLOs with your peers" with a mean score of 6.94 for those principals meeting their student achievement goals versus 5.63 for those who did not is worth further investigation, as it demonstrates that principals share data when they are positive, but may have a reluctance to share data that are not positive with their peers.

When this is compared to the data for SLO 2, culture and climate, the mean scores follow a more assumed pattern, with principals meeting their SLO (*n* of 20) sharing their data and lessons learned with the peers who did not (*n* of 6). See Figure 28.

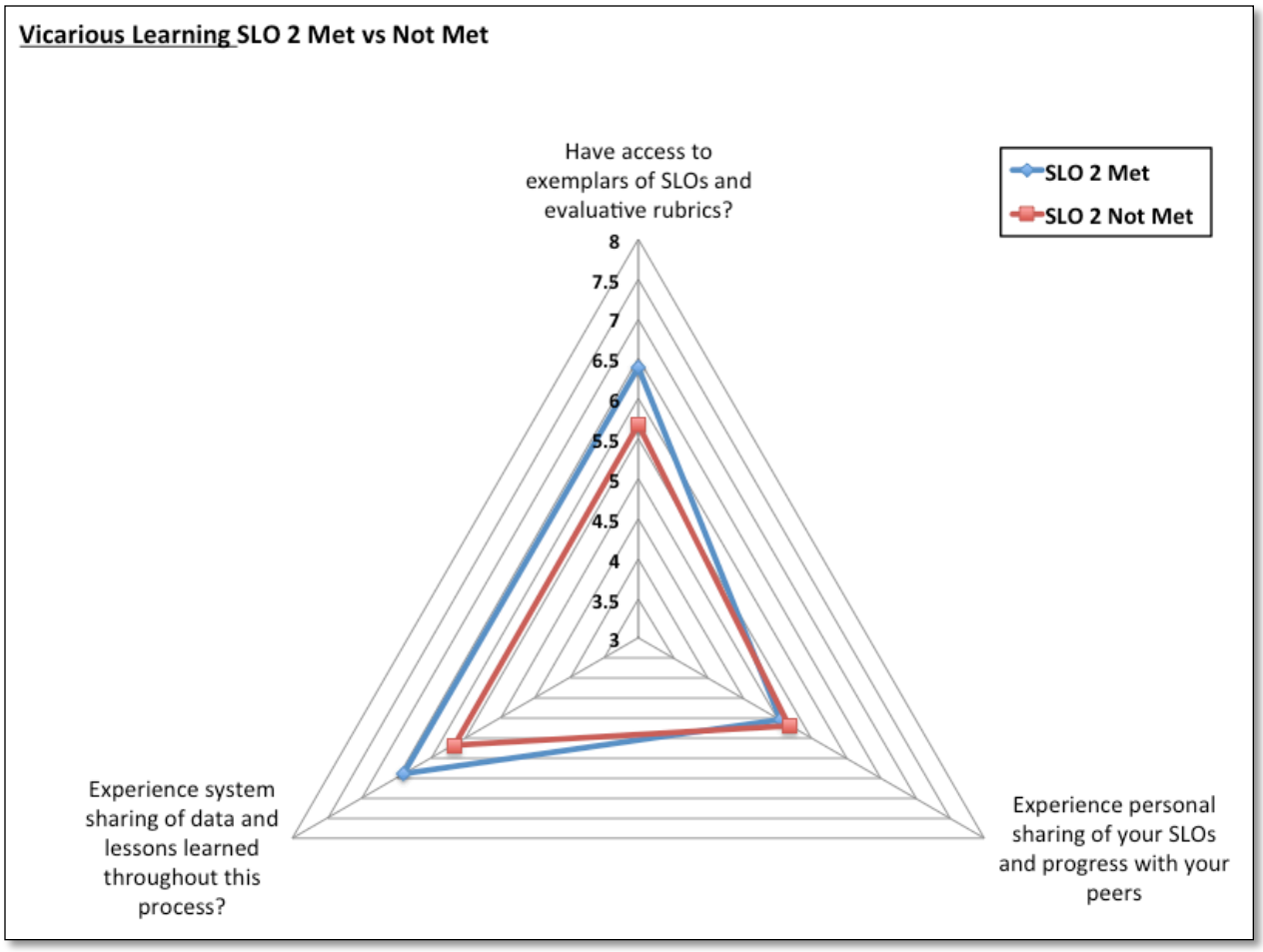

Figure 28. Radar graph, comparison of vicarious learning by SLO 2 status.
Note: Original graphic generated in Microsoft Excel.

For SLO 2, climate and culture, the mean scores for "experience personal sharing of your SLOs and progress with your peers" was much closer with a mean of 5.05 for those who met their SLO compared to a mean of 5.17 for those who did not. The other two questions support the hypothesis.

The third determinant of self-efficacy, verbal persuasion, Figure 29, follows the hypothesis that SLO attainment will yield a higher mean score with the question of "receive suggestions of alternative strategies to achieve your SLO" almost matching mean scores across both groups.

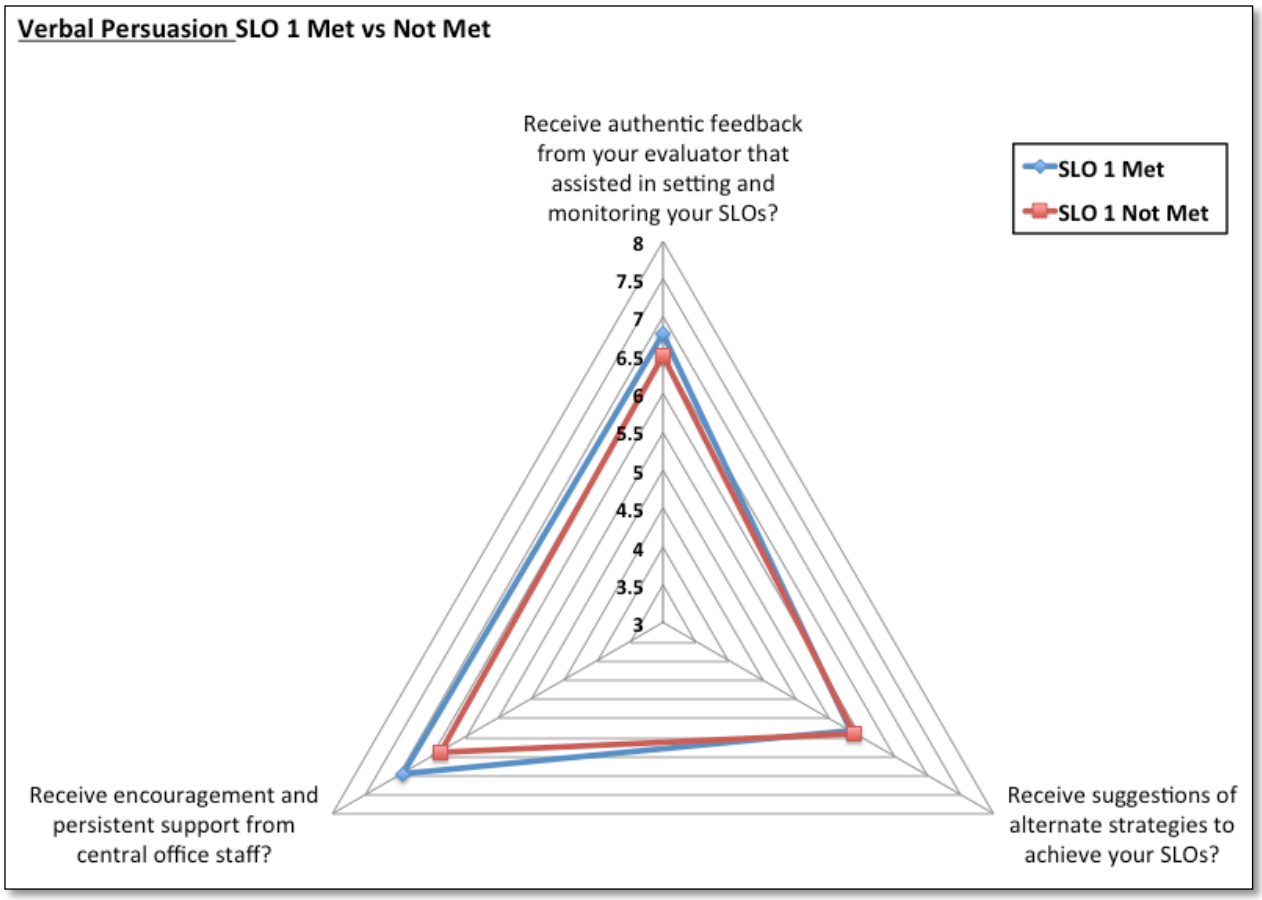

Figure 29. Radar graph, comparison of verbal persuasion by SLO 1 status.
Note: Original graphic generated in Microsoft Excel.

When examining the radar graph for verbal persuasion relative to SLO 2 attainment, Figure 30, those principals not meeting their SLO 2 showed a higher mean score for "receive encouragement and persistent support from central office staff" – mean score of 7.33 for those not meeting, compared to 6.6 for those who did. This would also make sense, as support from their evaluator should be more evident for those who may be challenged meeting their goals. This is also demonstrated in the mean scores for "receive authentic feedback from your evaluator that assisted in setting and monitoring your SLOs" with a mean score of 6.83 for those not meeting versus 6.65 for those who did. Again, the results can be interpreted as positive, for those who struggling should be

provide meaningful feedback throughout the process.

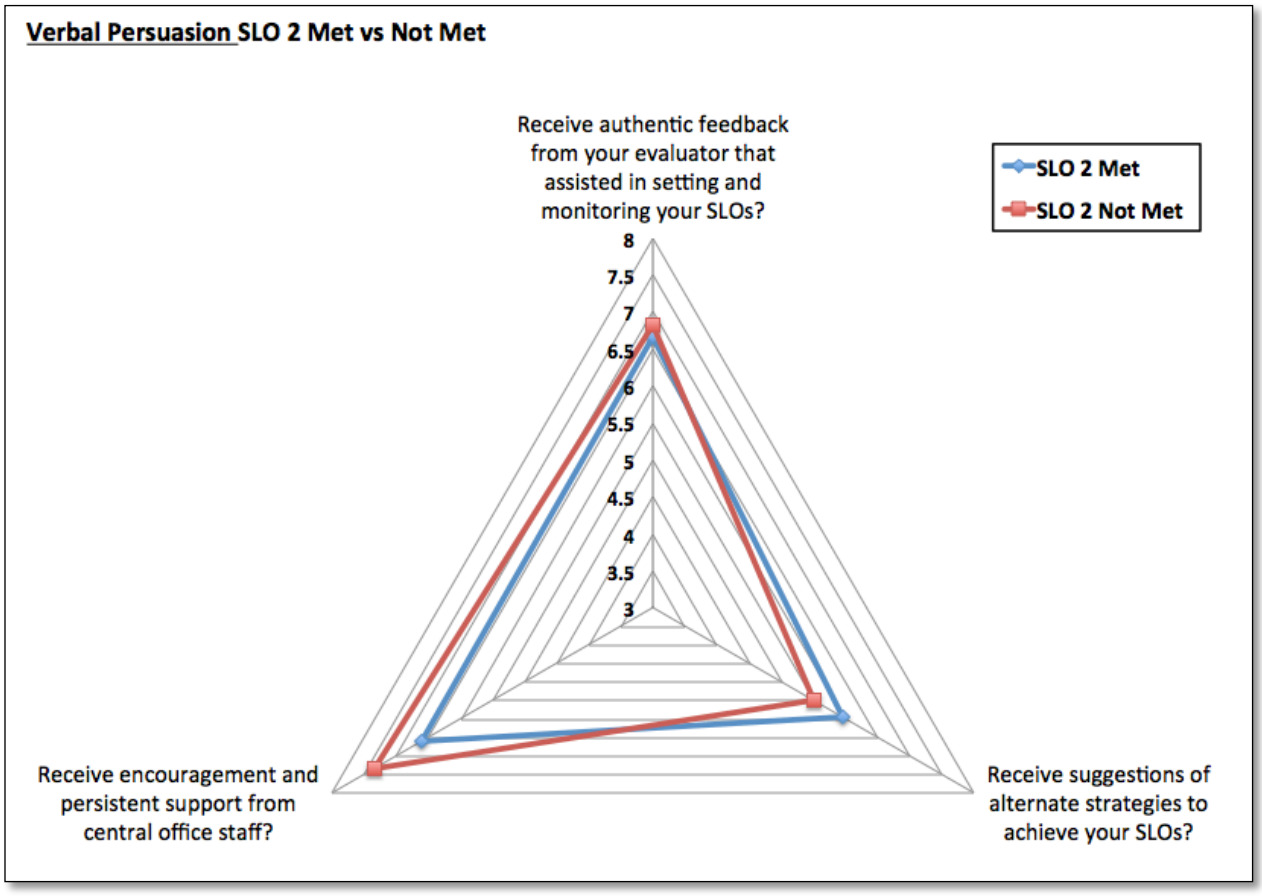

Figure 30. Radar graph, comparison of verbal persuasion by SLO 2 status.
Note: Original graphic generated in Microsoft Excel.

The final determinant of self-efficacy explored in the survey is physiological arousal. When examining the radar graph created for SLO 1, Figure 31, the data are as predicted with the exception of "perceive the SLOs as an authentic evaluation of your leadership." On this question, principals meeting their SLO 1 had a mean score of 5.72 compared to those who did not, with a mean score of six. This can be interpreted as a positive data point, as the respondents, even though they did not reach their target, are still identifying the process as an authentic measurement of their leadership.

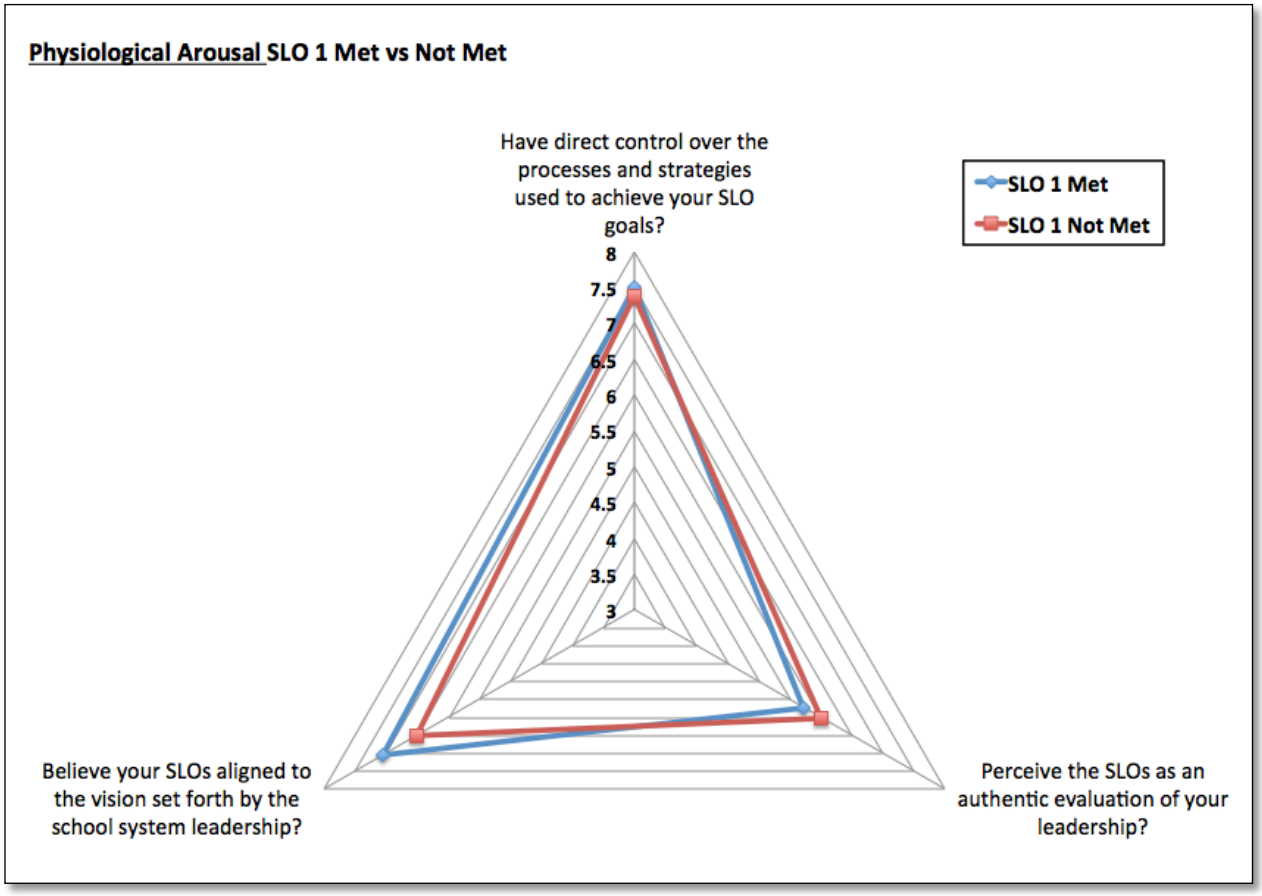

Figure 31. Radar graph, comparison of physiological arousal by SLO 1 status.
Note: Original graphic generated in Microsoft Excel

The final radar graph, Figure 32, reflects the data of SLO 2 attainment and physiological arousal. The data are the inverse of the hypothesis, with the highest mean scores attributed to those who did not meet their SLO for climate and culture. The results should be interpreted cautiously, as this represents a very small sample size of six principals who did not meet their SLO 2 compared to twenty who did. With that said, the

six responding principals posted the highest mean score of eight on the survey for the question "believe your SLOs aligned to the vision set forth by the school system leadership." This again may be interpreted as positive if the system is seeking alignment in effort with regard to climate and culture initiatives.

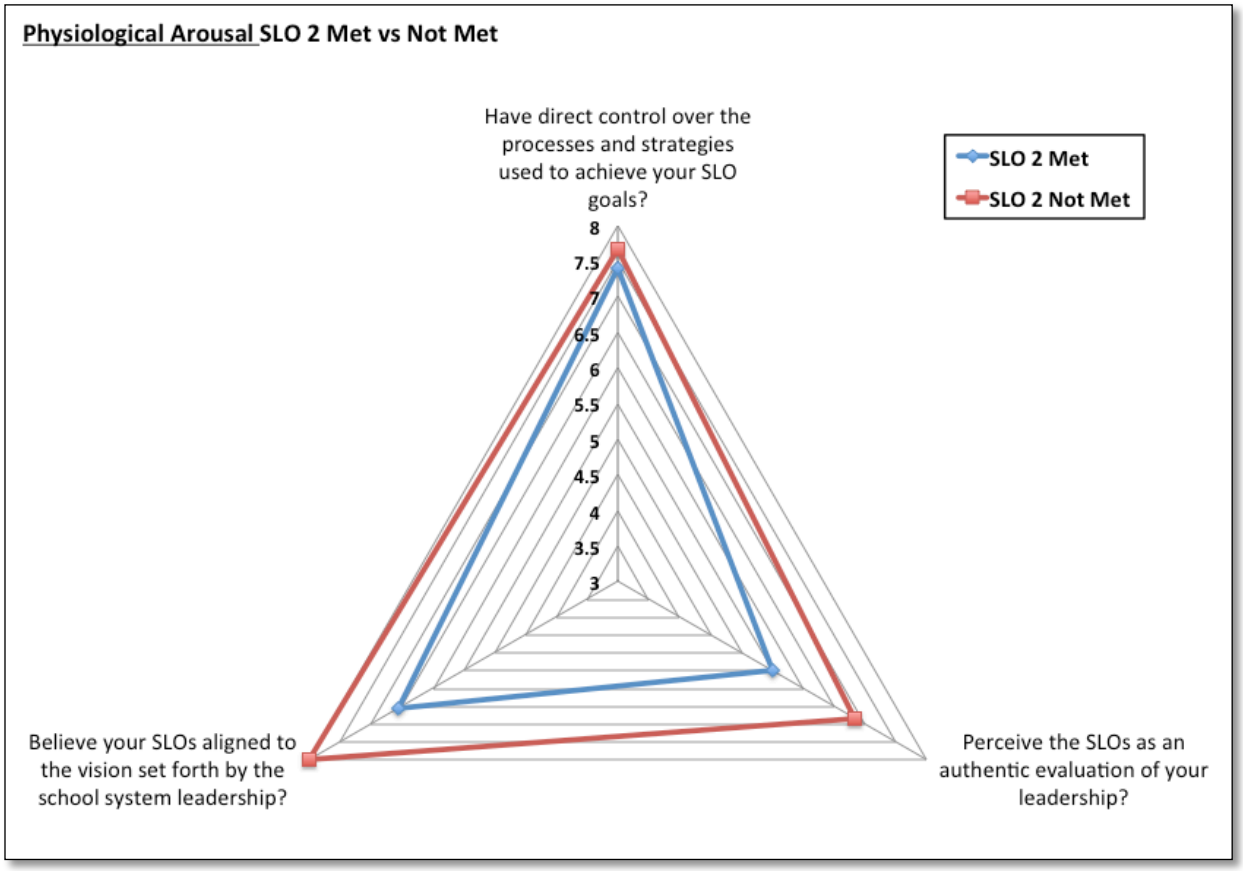

Figure 32. Radar graph, comparison of physiological arousal by SLO 2 status.
Note: Original graphic generated in Microsoft Excel.

Conclusions

This study examined the self-efficacy beliefs of the principals in a mid-size, Mid-Atlantic school system and how the attainment of the SLOs they selected for student achievement and culture and climate may have affected their responses. It also explored whether the four determinants of self-efficacy, as defined by Albert Bandura, were evident in their evaluation process and if there was statistical significance between those principals who met their SLOs and those who did not. The conceptual framework for the study was based on utilizing a proven survey tool of principal self-efficacy, the PSES developed by Tschannan-Moran and Gareis, and applying the variable of attainment of each principal's SLOs to results. The second part of the study was rooted in the research of Albert Bandura and his defined sources of self-efficacy: enactive mastery learning, vicarious learning, verbal persuasion, and physiological arousal. Earlier research has shown that principals with a higher sense of self-efficacy are more willing to attempt tasks that they perceive as challenging and that once they encounter challenge they will have a great likelihood to persevere in their pursuits. Considering the importance of the leadership of principals and the emphasis placed on quantifying their impact on student achievement and on the climate and culture of the schools they lead, understanding how best to support them in this work is essential.

As discussed in Section I of this study, a report co-authored by the NAESP and the NASSP found that effective principal evaluation systems are based on seven beliefs: that they be created by and for principals; are part of a comprehensive system of support and professional development; are flexible enough to accommodate differences in principals' experiences; are relevant to the improvement of principals' dynamic work; are

based on accurate, valid, and reliable information gathered through multiple measures; are fair in placing a priority on outcomes that principals can control; and useful for informing principals' learning and progress (Clifford, M., & Ross, 2013).

Related research by Oksana Parylo in 2012 observed that "roles and responsibilities have gone through considerable changes in the accountability era" and that those who had been principals for more than ten years had seen their evaluation evolve from "more of a managerial type, where you were evaluated on how well you managed people, dealt with a budget and dealt with the community" (p. 223). Evaluation systems now are more data driven and performance based. To balance this, principals valued evaluation systems that were a process—not an event—built on transparency, valuing dialogue, predicated on trust and respect, and underpinned with feedback and support (Parylo, 2012).

Trust and support was also a recurring theme in the research done by the National Comprehensive Center for Teacher Quality (Clifford, M., 2012). That report concluded that new evaluation systems should not only hold principals accountable for performance, but must also generate "trust among stakeholders" and "support principals' continued growth" (p. 59).

At the core, the most effective principal evaluation systems are those that have been created with direct input from the principals and reflect the reality of their daily job (Catano, 2006). They explicitly state the expectations and the manner in which they will be measured (Clifford, 2012; Lipscomb, 2012; Parylo, 2012). They are flexible tools that evolve as the role of the principal responds to the needs of the school and staff (Casserly, 2013). Principal evaluation systems must be relevant to the work being done by the

principal and must incorporate the standards of practice that are widely accepted (Browns-Simms, 2010). The goal of evaluation is to improve performance; this is done by selecting relevant measures that reflect the reality of a principal's job (Parylo, 2012; Clifford, H., 2013; Clifford H., 2012). Evaluation instruments must inform professional development, as the cycle of evaluation is the setting of expectations and goals (Canole, 2013). Evaluation systems must be transparent and systematically applied to all principals with the highest priority placed on outcomes the principals control (Clifford, M., 2013). There must be no "surprises"; if there is an area of need, it must be identified in a manner that can be addressed and corrected (Parylo, 2012). There is considerable evidence that principal evaluation must be collaboratively developed, reflective of the work of the principal, and target driven.

Principals must have the confidence in the decisions they make as the instructional leader of the school (Tschannen, 2005). Bandura (2000) explained that, "when faced with obstacles, setbacks, and failures, those who doubt their capabilities slacken their efforts, give up, or settle for mediocre solutions. Those who have a strong belief in their capabilities redouble their efforts to master the challenge" (p. 120). This is an essential element of effective SLO goal setting: a principal's belief in his or her own abilities. Combining social cognitive theory and self-efficacy, McCormick (2001) states, "Successful leadership involves using social influence processes to organize, direct, and motivate the actions of others. It requires persistent task-directed effort, effective task strategies, and the artful application of various conceptual, technical, and interpersonal skills" (p. 28).

SLOs and their inclusion in a principal evaluation system address and incorporate

many of the recommended elements of principal evaluation. They are flexible, self-selected, authentic, and directly related to the work of the principal. Moreover, the process by which they are developed, monitored, modified, and incorporated into a significant portion of a principal's annual evaluation give them weight and focus the work of the principal. While this study did not gather data on the individual SLOs, responses from the principals on the second section of the survey indicate that they are an institutional element of the evaluation system in this school system examined. Further study into the SLOs themselves would be beneficial.

While this study did not demonstrate a statistically significant relationship between attainment of SLOs and the overall mean scores, it did explore self-efficacy perceptions of the principals and deliver results consistent with previous applications of the PSES. Data collected at the sublevel construct for the PSES—instructional leadership, moral leadership, and managerial leadership—show that, regardless of the tool being used to evaluate principals, these are foundational elements of a principal's work and are worthy of examination.

Moreover, the second section of the survey demonstrated that the participants encountered elements of Albert Badura's determinants of self-efficacy throughout the process and reinforce that further exploration into this topic may yield insightful data. Embedded within the questions were concepts of agency, such as question 28 *"Throughout the evaluation process to what extent did you have direct control over the processes and strategies used to achieve your SLO goals*," and authenticity, such as question 29 "*Throughout the evaluation process to what extent did you perceive the SLOs as an authentic evaluation of your leadership?"* Each of these elements is considered to

be an essential component in an effective evaluation system. Ultimately, only so much can be gleaned from an ordinal response on a Likert scale, but the data do drive the conversation down a path supported by recent research.

Finally, limitations of sample size must also be considered, as the data set for this was only twenty-six principals in one school system. Principals throughout the state are all required to have SLOs as a significant portion of their annual evaluation. Future researchers could use surveys to gather data from across the state, which may yield a more significant statistical result.

Recommendations for Further Research

Future study should be done with a larger sample size, perhaps across the state, and if done on a larger scale, gathering respondent demographic, educational, and assignment information would be recommended. This would provide another way of cross-tabbing the data that may yield statistically significant results. This was not done for this study, as the sample size was so limited that this information would have impacted the anonymity of participants.

Beyond demographic data, gathering more information about the SLOs selected by the principals beyond a simple "yes" or "no" on attainment is also recommended. Additional information on the SLOs could be incorporated into the survey and would allow more detailed analysis of results. SLOs data could be categorized into related sets so that further analysis could be done with regard to those meeting and not meeting their identified outcomes. For example, principals selecting specific State PARCC scores could be grouped together based on the content and grade level of students in the identified SLO cohort. For culture and climate SLOs, principals identifying similar

outcomes could be compared against system-wide data for like cohorts in other schools. This could also inform evaluators so they may be better able to support principals in determining SLOs that are challenging yet attainable and specifically tailored to the needs of their particular assignment.

Once the SLOs data have been incorporated, having the principals take the survey and running similar correlative tests would provide results that could be applied to selecting participants for follow-up interviews to explain anomalies, such as those seen for SLO 2 and moral leadership, where the principals in this study had higher mean scores for those not attaining their SLOs versus those who did. These data would be beneficial in refining the second set of questions based on Bandura's determinants of self-efficacy, as the questions were created for this survey and have not been psychometrically validated.

The ultimate goal would be to unpack the extent to which Bandura's determinants of self-efficacy are embedded in the evaluation process and if administered over successive years, longitudinal data could be examined to better inform the practice of the evaluators of principals. It is this second set of data on Bandura's determinants that may have the most relevant use for the school system.

Implications for the School System

The data gathered in the second section of the survey reflect the work of the principal's evaluator, specifically, the deputy superintendent. The role of the deputy superintendent is to work directly with the principal at the beginning of the school year to review the previous year's data and select a SLO for student achievement and one for culture and climate. This work requires the principal and the deputy superintendent to

have a comprehensive understanding of the performance of the school for both achievement, as well as behavioral data. The deputy superintendent must work with the principal to set goals that are rigorous, but also attainable. The goals must also align with the work of the school and the system. The extent to which the principal responses indicated that they are provided opportunities to demonstrate incremental mastery or share their experiences and knowledge show the engagement of the evaluator with those they are evaluating. How the evaluator supported and encouraged the principal by providing meaningful, authentic feedback could be measured against the principal's perceptions and professional growth. Data gathered on SLO attainment could be compared against the responses provided by the principal and this, in turn, could aid the evaluator in better understanding what assisted the principal in meeting their SLO. No tool to provide feedback to the evaluator of principals about the impact of their work has yet been developed; the responses from the second section of the survey could be an informative place to start.

Finally, the questions asked of the principals with regard to their perceptions of how aligned their work is with the work of the school system, how much control they exert over the process and the results, as well as how authentic and valuable the process is, have implications for the system as a whole. The data gathered in the second section of this survey could be used to measure the extent to which system initiatives regarding student achievement and culture and climate are being communicated to principals and incorporated into the professional goals they set for themselves. It could also measure the extent to which the principals believe they are personally involved in system-level initiatives, such as question 12, "*To what extent do you shape the operational policies*

and procedures that are necessary to manage your school?" and question 32, *"Throughout the evaluation process to what extent do you believe your SLOs aligned to the vision set forth by the school system leadership?*" The data gathered from these questions, correlated to the respondent's attainment of their SLO, could be examined to determine how successful system work has been at aligning efforts and delivering meaningful results, both on a system scale and on a personal level.

A great deal of effort has gone into principal and teacher evaluation in Maryland over the past decade, with much of it compelled by legislation and federal mandate. All of it has been driven by the need to improve public education and increase student achievement. School principals are at the center of this work, and understanding how best to support them and their professional growth must always be at the center of evaluation. Principals' belief in their self-efficacy, in their ability to lead others instructionally, morally, and managerially, must be reinforced and advanced. Understanding the sources of self-efficacy and building these into evaluation systems that incorporate enactive mastery, vicarious learning, verbal persuasion, and physiological arousal is essential if we are to grow our leaders and help them, in turn, grow the self-efficacy of their teachers and their students.

Appendices

Appendix A: Survey Instrument

Principal Questionnaire

This questionnaire is designed to help us gain a better understanding of the kinds of things that create challenges for principals in their school activities.

Directions: Please indicate your opinion about each of the questions below by marking one of the nine responses in the columns on the right side. The scale of responses range from "None at All" (1) to "A Great Deal" (9), with "Some Degree" (5) representing the midpoint between these low and high extremes. Your answers are confidential.

Please respond to each of the questions by considering the combination of your current ability, resources, and opportunity to do each of the following in your present position.

"In your current role as principal, to what extent can you…"

1. facilitate student learning in your school?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

2. generate enthusiasm for a shared vision for your school?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

3. handle the time demands of your job?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

4. manage change in your school?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

5. promote school spirit among a large majority of the student population?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

6. create a positive learning environment in your school?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

7. raise student achievement on standardized tests?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

8. promote a positive image of your school with the media?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

9. motivate teachers?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

10. promote the prevailing values of the community in your school?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

11. maintain control of your own daily schedule?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

12. shape the operational policies and procedures that are necessary to manage your school?

Not at		Very		Some		Quite a		A Great

All		Little		Degree		bit		Deal
1	2	3	4	5	6	7	8	9

13. handle effectively the discipline of students in your school?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

14. promote acceptable behavior among students?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

15. handle the paperwork required of the job?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

16. promote ethical behavior among school personnel?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

17. cope with the stress of the job?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

18. prioritize among competing demands of the job?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

While working through the SLO process, to what extent did you:

19. modify targets based on formative data?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

20. document and celebrate incremental successes throughout the process?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

21. document evidence of support and intervention for your SLOs?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

22. attain the student achievement targets set as your SLO?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

23. attain the program/school-wide targets set as your SLOs?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

24. have access to exemplars of SLOs and evaluative rubrics?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

25. experience personal sharing of your SLOs and progress with your peers?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

26. experience system sharing of data and lessons learned throughout this process?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

27. receive authentic feedback from your evaluator that assisted in setting and monitoring your SLOs?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

28. receive suggestions of alternate strategies to achieve your SLOs?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

29. receive encouragement and persistent support from central office staff?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

30. have direct control over the processes and strategies used to achieve your SLO goals?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

31. perceive the SLOs as an authentic evaluation of your leadership?

Not at All		Very Little		Strong Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

32. believe your SLOs aligned to the vision set forth by the school system leadership?

Not at All		Very Little		Some Degree		Quite a bit		A Great Deal
1	2	3	4	5	6	7	8	9

Appendix B: Survey Recruitment Email to SMCPS Principals

Re: Examining Principal Self-Efficacy and the SLO process
From: _____, Administrative Accountability Officer
To: Principals

Dear Principal,

Our colleague, Scott Smith, has asked our assistance in collecting data about Student Learning Outcomes and Principal Self-Efficacy Beliefs. You are being recruited to participate because you have recently gone through the evaluation and SLO process. The information gathered from this survey will be able to help SMCPS in strengthening principal evaluation, specifically the SLO process.

His survey is being sent to all principals and is part of his dissertation conducted under the direction of Dr. Patricia Richardson at the University of Maryland. The study has been approved by the Institutional Review Board (IRB) of the University of Maryland and the SMCPS Office of Strategic Planning.

This study will be conducted through a brief 10-minute survey that you will be able to access electronically through the University of Maryland's Qualtrics survey platform link September 1–14, 2017: (insert link here). When you click on the link, you will find a "Letter of Consent." If you choose to participate in his survey, click "Yes" and the survey will open for you. If you do not choose to participate, please click "No" and the survey will close.

Our goal is to have 100% of our principals complete the survey. As an incentive, Mr. Smith is offering $5.00 Starbucks gift cards to all participants. Please help him reach his goal of 100% participation!

I encourage you to participate. Please note your employment status in the county will not be affected by your participation or non-participation in this study.

Thank you,

_____, Administrative Accountability Officer

Appendix C: IRB Approval Letter

UNIVERSITY OF MARYLAND
INSTITUTIONAL REVIEW BOARD

1204 Marie Mount Hall
College Park, MD 20742-5125
TEL 301.405.4212
FAX 301.314.1475
irb@umd.edu
www.umresearch.umd.edu/IRB

DATE:	August 15, 2017
TO:	James Smith, MA
FROM:	University of Maryland College Park (UMCP) IRB
PROJECT TITLE:	[1104486-1] Student Learning Outcomes and Their Relationship to Principals' Self-Efficacy
REFERENCE #:	
SUBMISSION TYPE:	New Project
ACTION:	APPROVED
APPROVAL DATE:	August 15, 2017
EXPIRATION DATE:	August 14, 2018
REVIEW TYPE:	Expedited Review
REVIEW CATEGORY:	Expedited review category # 7

Thank you for your submission of New Project materials for this project. The University of Maryland College Park (UMCP) IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit ratio and a project design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

Prior to submission to the IRB Office, this project received scientific review from the departmental IRB Liaison.

This submission has received Expedited Review based on the applicable federal regulations.

This project has been determined to be a Minimal Risk project. Based on the risks, this project requires continuing review by this committee on an annual basis. Please use the appropriate forms for this procedure. Your documentation for continuing review must be received with sufficient time for review and continued approval before the expiration date of August 14, 2018.

Please remember that informed consent is a process beginning with a description of the project and insurance of participant understanding followed by a signed consent form. Informed consent must continue throughout the project via a dialogue between the researcher and research participant. Unless a consent waiver or alteration has been approved, Federal regulations require that each participant receives a copy of the consent document.

Please note that any revision to previously approved materials must be approved by this committee prior to initiation. Please use the appropriate revision forms for this procedure.

All UNANTICIPATED PROBLEMS involving risks to subjects or others (UPIRSOs) and SERIOUS and UNEXPECTED adverse events must be reported promptly to this office. Please use the appropriate reporting forms for this procedure. All FDA and sponsor reporting requirements should also be followed.

All NON-COMPLIANCE issues or COMPLAINTS regarding this project must be reported promptly to this office.

References

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2)*,* 191–215.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117–148.

Bandura, A. (2000) Cultivate self-efficacy for personal and organizational effectiveness. In Edwin A. Locke (Ed.) The Blackwell Handbook of Principles of Organizational Behavior (pp. 120-136). Malden, MA: Oxford University Press.

Bandura, A., & Locke, E.A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology, 88*(1), 87–99.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York, NY: W. H. Freeman and Company

Branch, G.F., Hanushek, E.A., & Rivkin, S.G. (2013). School leaders matter. *Education Next, 13*(1), 62–69.

Brown-Sims, M., & National Comprehensive Center for Teacher Quality. (2010). Evaluating school principals. tips & tools. *National Comprehensive Center for Teacher Quality.*

Casserly, M., Lewis, S., Simon, C., Uzzell, R., Palacios, M., & Council of the Great City Schools. (2013). Principal evaluations and the principal supervisor: Survey results from the great city schools. *Council of the Great City Schools*.

Clark, J.K., Evans, A.T., & Wegener, D.T. (2011). Perceptions of source efficacy and persuasion: Multiple mechanisms for source effects on attitudes. *European Journal of Social Psychology, 41*(5)*,* 596–607.

Clifford, M., Hansen, U. J., Wraight, S., & National Comprehensive Center for Teacher Quality. (2012). A practical guide to designing comprehensive principal evaluation systems: A tool to assist in the development of principal evaluation systems. *National Comprehensive Center for Teacher Quality*.

Clifford, M., Menon, R., Gangi, T., Condon, C., Hornung, K., & American Institutes, f. R. (2012). Measuring school climate for gauging principal performance: A review of the validity and reliability of publicly accessible measures. A quality school leadership issue brief. *American Institutes for Research*.

Clifford, M., & Ross, S. (2013). Rethinking Principal Evaluation. *National Association of Secondary School Principals and the National Association of Elementary Principals.* Alexandria, VA.

Connolly, M. (2007). Harried principals aren't helpful principals. *Principal, 86*(5), 33–35.

Council of Chief State School Officers. (2015) ISLLC 2015: Model Policy Standards for Educational Leaders. Washington, D.C.

Drago-Severson, E., Blum-DeStefano, J., & Asghar, A. (2013). *Learning for leadership: Developmental strategies for building capacity in our schools.* Thousand Oaks, CA: Corwin.

Fisher, Y. (2014). The timeline of self-efficacy: Changes during the professional life cycle of school principals. *Journal of Educational Administration, 52*(1), 58–83.

Federici, R. A. (2012). Principals' self-efficacy: Relations with job autonomy, job satisfaction, and contextual constraints. *European Journal of Psychology of Education*.

Friedman, I.A. (2002). Burnout in school principals: Role related antecedents. *Social Psychology of Education, 5*(3)*,* 229–251.

Gaziel, H. (2008). Principals' performance assessment: Empirical evidence from an Israeli case study. *Educational Management Administration & Leadership, 36*(3), 337–351.

Goddard, R.D., Hoy, W.K., & Woolfolk Hoy, A. (2004). Collective efficacy beliefs: Theoretical developments, empirical evidence, and future directions. *Educational Researcher, 33*(3), 3–13.

Goldring, E., Porter, A., Murphy, J., Elliott, S. N., & Cravens, X. (2007). Assessing Learning-Centered Leadership: Connections to Research, Professional Standards, and Current Practices*. Vanderbilt University, Learning Sciences Institute*. Nashville, TN.

Jackson, J.W. (2002). Enhancing self-efficacy and learning performance. *The Journal of Experimental Education, 70*(3), 243–254.

Lipscomb, S., Chiang, H., Gill, B., & Mathematica, P. R. (2012). Value-added estimates for phase 1 of the Pennsylvania teacher and principal evaluation pilot. *Mathematica Policy Research, Inc*.

Maryland State Department of Education. (2013) *The Maryland Teacher and Principal Evaluation Guidebook.* Baltimore, MD.

Margolis, H., & McCabe, P.P. (2006). Improving self-efficacy and motivation: What to do, what to say. *Intervention in School and Clinic, 41*(4), 218–227.

Maurer, T. J., & Andrews, K. D. (200). Traditional, Likert, and Simplified Measures of Self-Efficacy. Education and Psychological Measurement, Vol. 60 No. 7, December 2000, 965–973.

McCollum, D.L., & Kajs, L.T. (2009). Examining the relationship between school administrators' efficacy and goal orientations. *Education Research Quarterly, 32*(3), 29–46.

McCormick, M. J. (2001). Self-efficacy and leadership effectiveness: Applying social cognitive theory to leadership. Journal of Leadership Studies, 8(1), 22–33.

Mendels, P. (2012). Principals in the pipeline: Districts construct a framework to develop school leadership. *Journal of Staff Development, 33*(3), 48–52.

Norton, M. S. (2002). Let's keep our quality school principals on the job. *The High School Journal, 86*(2), 50–56.

Pajares (2002). Overview of social cognitive theory and of self efficacy. Retrieved July 13, 2017 from http://www.emory.edu/EDUCATION/mfp/eff.html.

Parylo, O., Zepeda, S. J., & Bengtson, E. (2012). Principals' experiences of being evaluated: A phenomenological study. *Educational Assessment, Evaluation and Accountability, 24*(3), 215–238.

Porter, A. C., Polikoff, M. S., Goldring, E., Murphy, J., Elliott, S. N., & May, H. (2010). Developing a psychometrically sound assessment of school leadership: The VAL-ED as a case study. *Educational Administration Quarterly, 46*(2), 135–173.

Rowan, B., & Denk, C. E. (1984). Management succession, school socioeconomic status, and basic skills achievement. American Educational Research Journal, 21(3), 517–537.

Schyns, B. (2004). The influence of occupational self-efficacy on the relationship of leadership behavior and preparedness for occupational change. *Journal of Career Development, 30*(4), 247–261.

Simon, Mark 1, 2, M.Simon@epi.org. (2012). A tale of two districts. *Educational Leadership, 70*(3), 58–63.

Slotnik, W.J., Bugler, D., & Liang, G. (2014) Real Progress in Maryland, Student Learning Objectives and Teacher and Principals Evaluation. *Mid-Atlantic Comprehensive Center.* Washington, D.C.

Slotnik, W.J., Bugler, D., & Liang, G. (2015) Change in Practice in Maryland, Student Learning Objectives and Teacher and Principal Evaluation. *Mid-Atlantic Comprehensive Center*. Washington, D.C.

Slotnik, W.J., Bugler, D., & Liang, G. (2016) Stay the Course, Teacher and Principal Evaluation in Maryland. *Mid-Atlantic Comprehensive Center*. Washington, D.C. St. Mary's County Public Schools (2010, 2012, 2013). Leadership Performance Assessment System. St. Mary's County, MD.

St. Mary's County Public Schools (2015). 2015–2016 Performance Assessment System. Leonardtown, MD. Retrieved from https://eval.smcps.org/home.aspx.

Thomas, C.N., & Kearney, K. (2010) Effective principal support: What will it take? *Leadership, 40*(2), 8–11.

Tschannen-Moran, M., & Gareis, C.R. (2004). Principals' sense of efficacy: Assessing a promising construct. *Journal of Educational Administration, 42*(5), 573–585.

Tschannen-Moran, M. & Gareis, C.R. (2007). Cultivating principals' self-efficacy: Supports that matter. *Journal of School Leadership, 17*(1), 89–114.

The Wallace Foundation. (2009). Assessing the Effectiveness of School Leaders. New York City, NY.